MASTER

Optimization of strategic supply chain planning

Molhoek, T.H.B.

Award date:
2018

Eindhoven, December 2017

University of Technology Eindhoven

Master thesis project

Optimization of strategic supply chain planning

by T.H.B. Molhoek
BSc Industrial Engineering & Management Science
Student identity number 0663799

University Supervisors:
prof. dr. ir. I.J.B.F. Adan
prof. dr. A.G. de Kok

Company Supervisor:
Tomasz Wolańczyk

Abstract

Nowadays, the globalization leads to more intense competition, this forces companies to optimize their performance in order to survive. In this master thesis project we try to optimize the strategical planning in supply chain management. The strategical planning consists in this case of choosing the right location to manufacture products. Our decisions are based on the cost and the lead time performance of the supply chain. This problem is classified as a supply chain network design problem. We use a technique called stochastic programming, which is branch of mathematical optimization. The stochastic programming approach allows for the inclusion of uncertainty related to input parameters. This allows us to evaluate candidate solutions over a range of plausible future scenarios. Solving a stochastic programming problem can be quite hard, but we use the sample average approximation algorithm to handle this. In our results we observe that there is a trade off between the cost and lead time performance. Furthermore, we conclude that the current supply chain network design in place is dominated by solutions that we found.

Keywords:

Supply chain network design; Stochastic programming; Sample average approximation

Management summary

Alpha* is considered to be the global industry leader for green indoor climate control. Currently, they have five factories in Europe which collaborate to produce all of their products. A number of years ago they decided to move assembly operations to a factory with competitive labor cost compared to others. Now, Alpha wants to analyze this decision and optimize the locations for assembly operations based on cost and lead time. The main research question of this master thesis project is:

How can Alpha create more value from their production locations for their complete product portfolio. Specifically, do they have to change some assembly operation locations within their supply chain?

To answer this question, we first need to analyze the current situation in terms of cost and lead time. The analysis is based on historical data related to cost and demand for products. Alpha produces a wide range of products and they use product families to categorize them. Our analysis is on product family level. Before the analysis starts, we select the product families for which there is enough useful data available, this resulted in selecting 79 of the 187 product families. This selection represents approximately 73% of the value of all products sold in the last 12 months.

We identify the relevant cost based on the residual cash flow. In our case the relevant costs are the operating expenses, transportation cost, tax cost and change in working capital. All of these are influenced by the decision for the assembly location of a product family. In addition, we discount the residual cash flow obtain the discounted residual cash flow. We use the weighted average cost of capital as the discount rate, because of Alpha's financing structure. Finally, we analyze the change in lead time related to the decision for the assembly location of a product family. We propose to use relative weighted lead time performance difference, where the weight is determined by the average current cost price of a product family. The lead time performance is an important performance measure for Alpha and supply chain management in general.

This problem can be classified as a supply chain network design problem and we use the stochastic programming approach to solve this problem. The stochastic programming method allows for the inclusion of uncertainty related to the input parameters, which reflects the various plausible future scenarios. In our model we consider the input parameters for material cost, variable machine cost, labor cost, transportation cost and future demand as random variables. First, we formulate a mixed integer linear programming problem with deterministic input parameters. This is the basis for the stochastic programming formulation. The mixed integer linear programming is a bi-objective problem, because we have the cost and lead time in the objective function. Hereafter, we consider some of the input parameters in the mixed integer linear programming problem as random variables to arrive at the stochastic programming formulation. Solving a

*This is a masked name. The real company name is hidden due to confidentiality reasons.

stochastic programming problem can be quite hard, to solve this issue we use the sample average approximation algorithm. This algorithm basically samples values from the random variables, but this algorithm assumes they follow a known probability distribution. To obtain the probability distributions for our random variables, we try to fit a theoretical probability distribution. If we fail to fit a theoretical probability distribution for a random variable, then we build an empirical distribution from the available data.

We solved our model for various problem instances, which reflects various future scenarios for the random variables and varying importance of each objective. The importance of the objectives are quantified by assigning weights to both objectives. The result of doing so is a set of candidate solutions. They define the locations for assembly operations for all selected product families. We show there exists a Pareto front which is a set of optimal solutions for varying objective weights. The two objectives are conflicting, if we have better lead time performance, then the cost increases and vice versa. From the Pareto front we conclude that the current supply chain network design in place is dominated by the solutions that we found. Furthermore, Alpha could operate with more or less the same cost, but improve their lead time performance by using our solutions. The difference between the current and the dominating configuration in present value of the total cost is approximately 0.6%. Let us remark that this difference is in favor of the dominating configuration, i.e. lower costs. In addition, we did not calculate how much money Alpha gains from the improvement in lead time performance. Therefore, our conclusion is somewhat limited, but we showed that our proposed conceptual model works for this case study. The conceptual model should be further extended by including the lead time impact on finished goods inventory levels. With this extension one could calculate the impact on the total cost of the complete supply chain.

All of our conclusions and results are based on the quantitative model that we have build. These models are always a simplified representation of the real world. Although we verified and validated our model in this project, we want to highlight that all results should be interpreted with caution. Alpha should use the outcomes as a part of their decision making process and not solely use the information from our model. In addition, we observe that there is clear link between lead time and finished goods inventory levels. We set a scope for our project and we did not include the impact on finished goods inventory. This impact can however be significant and we recommend to do a thorough analysis to asses this impact.

Acknowledgements

At this moment when I'm writing this, my period as a student has almost come to an end. I had a great time at the University in Eindhoven. Although I was not in Eindhoven for the last year, because I did an exchange program in Hong Kong and thereafter this project in Poland. The last year was intense, but brought me great things, I met great people and learned working in different cultural environments. I would like to thank Alpha for giving me the great opportunity to do this project in their organization.

Throughout this project I got a lot of support from various people at the company and from the university. First, I would like to thank my mentor Ivo Adan for giving me the opportunity to do this project abroad. In addition, he also helped by critically reviewing my work and providing clear and meaningful feedback. I could not have completed this project without his help.

The project was at Alpha's factory in Poland and I had a very good time there. I really appreciated the whole global supply chain team, which I was part of during my project. The team made me feel welcome the first time I stepped in and it was great experience working with all the team members. I would especially like to thank my company mentor Tomasz Wolańczyk. He made me feel comfortable from the first I was in Poland. He did not only support with me with my thesis project, but he also helped me during the whole period I was living in Poland.

Lastly I would like to thank my girlfriend and my family for all the support that they gave me throughout my academic journey. Thank you Naomi, Mariëtte, Paul, Michiel for your support. I feel privileged with having such caring people around me who always provide with support when needed.

Preface

This document contains my master thesis, which I wrote in the partial fulfillment of the MSc degree in Operations Management & Logistics. It describes a case study for a supply chain network design problem. The name of the company will remain undisclosed due to confidentiality reasons. We use the name Alpha as masked name for the company.

Contents

Abstract i

Management summary ii

Acknowledgements iv

Preface v

1 Introduction 1
1.1 Alpha . . . 1
1.2 Problem context . . . 3

2 Research design 4
2.1 Research questions . . . 4
2.2 Research methodology . . . 6
2.3 Deliverables . . . 8
2.4 Project scope . . . 8

3 Supply chain network design 9
3.1 Facility location models . . . 9
3.2 Supply chain management aspects . . . 10
3.3 Uncertainty . . . 11
3.4 Robustness . . . 12

4 Conceptual model 14
4.1 Finance and supply chain . . . 14
4.2 Discounted residual cash flow . . . 16
4.3 Weighted average cost of capital . . . 17
4.4 Working capital . . . 19
4.5 Lead time performance . . . 20
4.6 Stochastic programming . . . 20

5 Modeling 24
5.1 SCND formulation . . . 24
5.2 Algorithmic Strategy . . . 26

6 Case study 29
6.1 Introduction . . . 29
6.2 Supply chain cost . . . 30
6.3 Lead time performance . . . 31

6.4 Investment cost . . . 32
6.5 Data . . . 32

7 Numerical results **38**
7.1 Problem instances . . . 38
7.2 Solution properties . . . 40
7.3 Verification and validation . . . 42

8 Implementation **43**

9 Conclusions and recommendations **44**
9.1 Conclusions . . . 44
9.2 Recommendations . . . 45

References **46**

Appendices **48**

A Cullen and Frey graph **49**

List of Tables

1 List of abbreviations and definitions . . . x

4.1 Operating expenses . . . 15
4.2 WACC formula symbols . . . 18

5.1 Definition of symbols . . . 25

6.1 Operating expenses . . . 30
6.2 Goodness of fit statistics . . . 34
6.3 OLS analysis statistics . . . 36

7.1 Optimality gaps from SAA algorithm in percentages . . . 39
7.2 Current situation for assembly locations . . . 41
7.3 Assembly locations for different values of λ and sample size N . . . 42

List of Figures

1.1 Supply chain structure (schematic) 2
1.2 Annual revenue of product families 3

2.1 Research model by Mitroff et al. (1974) 6

4.1 Cash flows in supply chain 14
4.2 Cash flows in supply chain 15
4.3 Schematic cash flows in Alpha group 16

6.1 Schematic drawing of the production process (i, j and k respectively represent the components factory, assembly factory and hub) 29
6.2 Box plots for assembly cost of product family 33
6.3 Labor cost per unit over time 34
6.4 Goodness-of-fit graphs for Gamma distribution 35

7.1 Cumulative probability curves of the total cost for different values of λ and N=20 40
7.2 Cumulative probability curves of lead time performance difference for different values of λ and N=20 40
7.3 Pareto front from calculating the objective function with N'=1000 41

A.1 Cullen and Frey graph 49

List of abbreviations

Table 1: List of abbreviations and definitions

Abbreviation	*Definition*
ARIMA	Autoregressive integrated moving average
CAPM	Capital asset pricing model
CFLP	Capacitated facility location problem
DRCF	Discounted residual cash flow
EBIT	Earning before interest and taxes
FTSE	Financial Times Stock Exchange
MILP	Mixed integer linear program
MLE	Maximum likelihood estimation
NOPAT	Net operating profit
OLS	Ordinary least squares
OR	Operations research
RCF	Residual cash flow
SAA	Sample average approximation
SCM	Supply chain management
SCND	Supply chain network design
UFLP	Uncapacitated facility location problem
WACC	Weighted average cost of capital
WIP	Work in progress

Chapter 1

Introduction

1.1 Alpha

The organization of interest for this master thesis project is Alpha X, which is a division of the Alpha group. Alpha group is a global engineering group and a FTSE 250 company on the London stock exchange. The total group consists of two more divisions: Alpha Y and Alpha Z. Alpha X is considered to be the global industry leader for green indoor climate control. This project is in collaboration with Alpha X. In 2016 they had an annual revenue of £290 million with an operating profit of £51.9 million with 1,900 employees. Hereafter, we will refer to Alpha X with the term Alpha.

Alpha has five factories in Europe and one in the USA. The factories in Europe are named by Jupiter, Saturn, Uranus, Neptune and Mars in this document. All factories are part of their supply chain except for the factory in the USA, which operates autonomously. A collection of four brands is covered by Alpha. These brands are all complementary, and in combination they can offer complete control systems. Alpha has in total three finished goods warehouses, also known as hubs, through which demand all over the globe is served. This project focuses on improving performance of their supply chain. The Saturn factory does assembly work only and they are mainly supplied by the Jupiter and Uranus factory or by a third party supplier. In Jupiter and Uranus they have foundries which are used to produce all sorts of metal shapes. These metal shapes get machined and if necessary are treated to obtain a semi-finished product. In the current situation the bulk of these semi-finished parts are sent to Saturn for an assembly operation. In the Neptune and Mars factories they produce respectively low-volume special products and vessels. These factories are in terms of volume and value much smaller than the other three.

Figure 1.1 depicts a schematic drawing of a part of the supply chain of Alpha. The Mars and Neptune locations are not included, because these factories are relatively small compared to the others. However, we will also include the Neptune factory in our analysis research project. In Figure 1.1 on the left side are all third party suppliers which supply raw materials as well as parts for the assembly operations. The Jupiter and Uranus factories have a semi-finished goods factory, i.e. the foundries that produce metal shapes. Semi-finished goods and third party suppliers items are the input for the assembly operations. The assembly operations produce the finished goods and these are stored in the finished goods warehouse also known as hubs. From the hubs these products get shipped to the end customer.

Currently Alpha has over 9,000 unique items that they can produce and sell. These

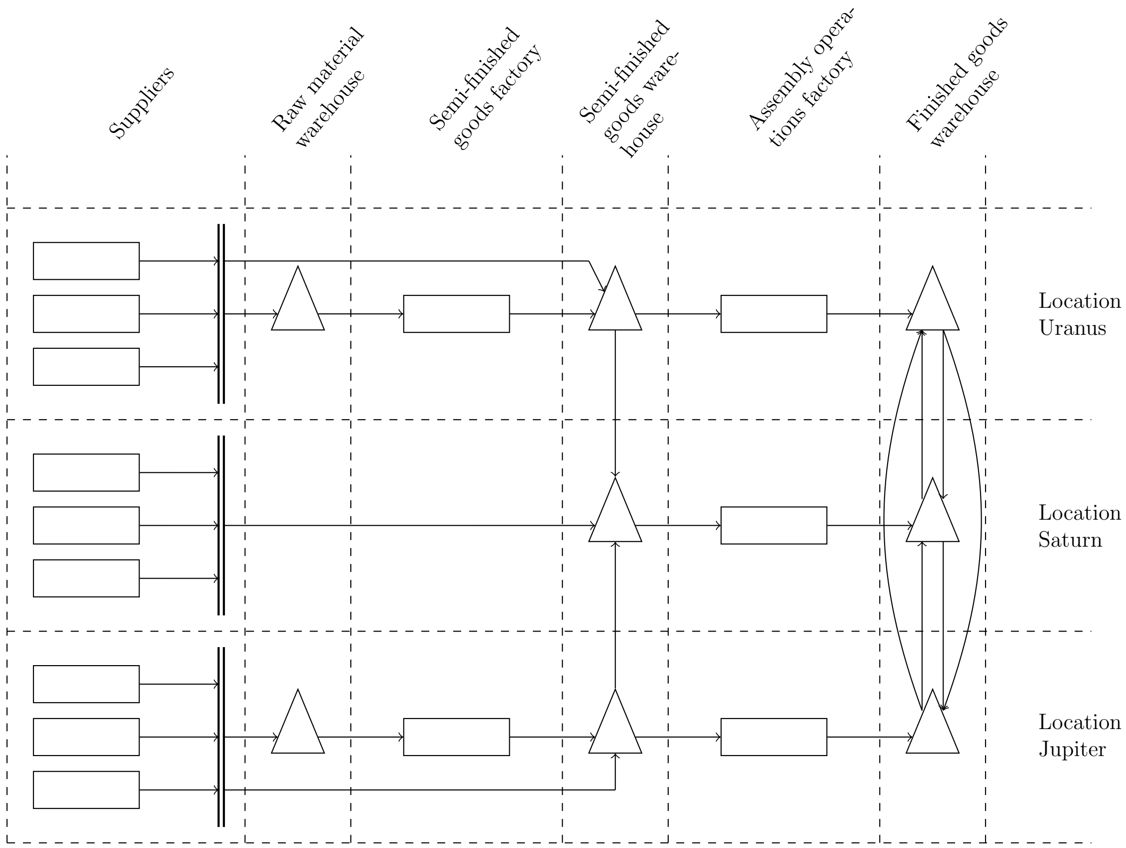

Figure 1.1: Supply chain structure (schematic)

unique items can be categorized into 187 product families. These are used to roughly plan the production capacity for these items. The product families all belong to a total of 7 categories. From analysis we see that 338 of these unique items are responsible for 80% of the total revenue in the last five years. These 338 items are from 50 different product families. This shows that Alpha has a huge portfolio of products with a lot of different products. Not all products are completely different, e.g. some can have small differences in packaging, colors or printing.

The scatter plot in Figure 1.2 depicts the mean annual revenue with respect to the coefficient of variation per product family. There seems to be a link between the mean revenue and the coefficient of variation. Higher revenues have a lower coefficient of variation than the lower revenues. Alpha can make accurate predictions on the value that will be sold for product categories, because sales teams get a target budget for these categories. This target budget is a total value that these teams need to sell in a fixed period, such as months, quarters or years. The forecasting on individual item level is difficult for the short term, i.e. on monthly basis. The demand planner mentions that the forecast accuracy on monthly basis is between 50% and 60% on individual item level. Although they can predict what total value is sold over a long term period, the mix of items in the total value can differ. Therefore, it makes sense to analyze different scenarios for the future on the mix of items in product families. The relocating of facilities is done on product family level.

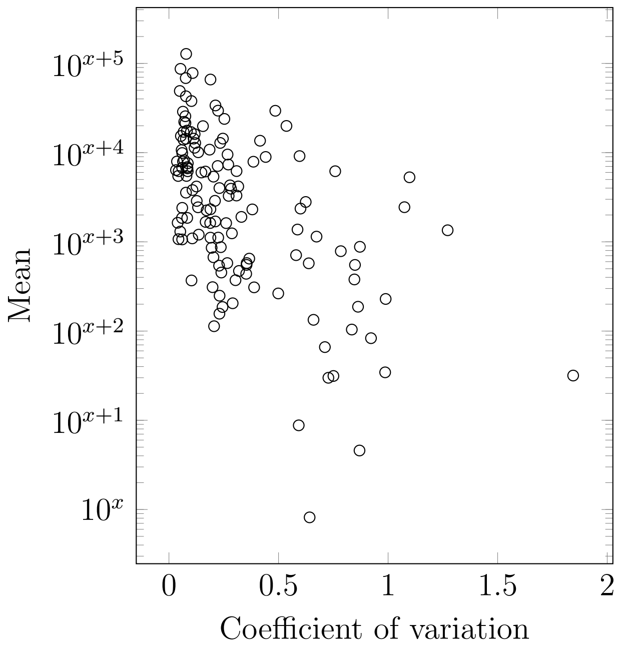

Figure 1.2: Annual revenue of product families

1.2 Problem context

An illustrative example of the current situation is a product which is produced in Jupiter and sent to Saturn for assembly. After assembly it is sent back from Saturn to Jupiter, because the customers for the product are located in Jupiter exclusively. This situation was created some years ago when Alpha decided to move a significant part of its assembly operations to Saturn. The main reason for this strategy was the low labor cost in the Saturn site compared to the other sites. Alpha wants analyze the efficiency of their supply chain in the current situation. Hereafter, they want to use this analysis to redesign their supply chain by making fact based decisions on assembly locations. The analysis and redesign should be based on four performance measures: cost, lead time and working capital. We use the weighted average cost of capital (WACC) as the discount rate to determine the present value of future costs. These metrics will ensure that the new design optimizes the supply chain network performance.

Chapter 2

Research design

We describe the research questions, research methodology, outline of this thesis, the scope of the project and the deliverables.

2.1 Research questions

Alpha wants to analyze their supply chain on four performance measures, namely: cost, lead time, working capital and WACC. This analysis should be focused on the flow of products through their supply chain. They want to know specifically what are the best locations to produce their products based on the performance measures. The possible locations are limited to the currently existing factory locations. Alpha wants to create more value from its supply chain operations. It means they should maximize the discounted residual cash flows (DRCF) from these supply chain operations. The exact definition of the DRCF is given in Chapter 4 on the conceptual model.

Main research question:

How can Alpha create more value from their production locations for their complete product portfolio. Specifically, do they have to change some assembly operation locations within their supply chain?

Current situation

The first phase in answering the main research question is analyzing the current situation. We want to know where they are selling their products. In addition, we want to know where these products are produced, what the lead time is and how much this costs. If we answer these questions, then we can make a first analysis of current situation at Alpha. Below is a list of detailed questions to be answered. The first question answers the question for the customers locations. The second question concerns the current lead time of their products and its flexibility. Questions three up to and including eight are to determine the current production location and costs associated with this location. In question 5 we use the term logistic cost. The last question in the list is to find out what the current level of working capital is. Questions nine, ten and eleven give information with respect to working capital, WACC and residual cash flow. These questions should provide an overview of the current situation based on the four performance measures.

1. *Where did Alpha sell their products?*
2. *What is the current lead time?*
3. *Where does Alpha currently manufacture their products?*
4. *What is the current production cost of every Alpha product?*
5. *What is the current logistic cost of every Alpha product?*
6. *Where does Alpha currently source their components from for every product?*
7. *What is the current cost of components for every Alpha product?*
8. *What are the logistic costs of components for every Alpha product?*
9. *What is the current level of working capital at Alpha?*
10. *What is the WACC for Alpha in the current situation?*
11. *What is current residual cash flow of Alpha from operations?*

Alternative situations

For the redesign we have a limited number of options. The possible alternative production locations are limited to the existing factory locations. To analyze these possible situations we want to answer the list of questions below. The first question is a sales forecast including the customers' locations. The second question describes the lead time for the alternative situations. Questions three up to and including six reveal the cost of producing on other locations. The seventh question tells what the required investment is for changing production locations. The remaining three questions answer respectively what the level of working capital, weighted average cost of capital and DRCF is for different production locations.

1. *Where will Alpha sell their products in the future and how much?*
2. *What would be the lead time if produced elsewhere?*
3. *What would be production costs elsewhere for every Alpha product?*
4. *What would be the logistics costs if produced elsewhere for every Alpha product?*
5. *What would be the cost of components for every Alpha product?*
6. *What would be the logistics costs of components if produced elsewhere for every Alpha product?*
7. *What investment would be required to produce at a different location?*
8. *What would be the level of working capital of Alpha when producing elsewhere?*
9. *What would be the WACC if Alpha produced elsewhere?*
10. *What would be the discounted residual cash flow at WACC of Alpha from operations when producing elsewhere?*

2.2 Research methodology

In the field of operations research (OR) it is common to use quantitative modeling to analyze real-world problems. This research project follows that approach and we use the framework described by Mitroff et al. (1974), see Figure 2.1. The frame work consists of four phases:

- Conceptualization
- Modeling
- Model solving
- Implementation

Mitroff et al. (1974) argue that a research cycle can start at any of these four phases. There exist different types of research projects that follow different routings in this model. In the next two paragraphs we describe types of quantitative operations research and their routing through the model depicted in Figure 2.1. Will M. Bertrand & Fransoo (2002)

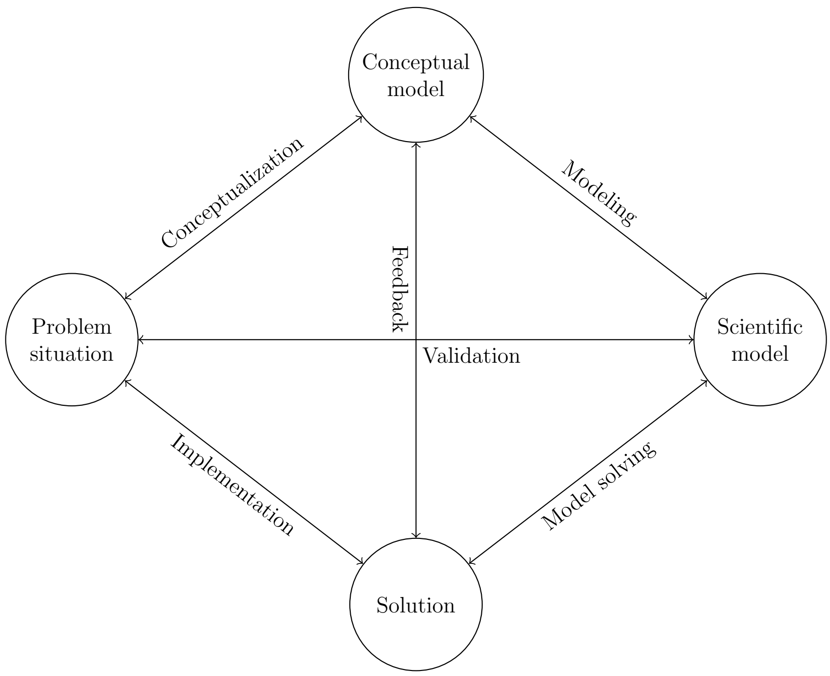

Figure 2.1: Research model by Mitroff et al. (1974)

describe a classification system for model-based OR research types. In addition, they link these to the research model of Mitroff et al. (1974). The classification is based on axiomatic or empirical and descriptive or normative research. Research projects do not exclusively belong to a certain classification, but they can belong to a combination. In axiomatic research the model plays a central role, sometimes it is even idealized. The focus is on obtaining solutions within the defined model that provide insight into the structure of the problem as defined by the model. Most of the research in the axiomatic domain is normative and some is descriptive. An example of an axiomatic descriptive type is queuing theory in which some “laws” are derived similar to the science of physics. This type of research typically only enters the modeling phase in the research model of

Mitroff et al. (1974). Inventory theory is an axiomatic normative type of research branch in operations research which has an extensive literature base. Normative axiomatic research usually follows the cycle modeling, model solving and feedback. The model solving process plays a central role in normative axiomatic research. Operations research uses formal methods from other branches of scientific research to obtain knowledge about these models. Mathematics, statistics and computer science deliver formal methods which are often used in the field of operations research.

Empirical descriptive researchers are interested in describing the causal relations between variables in a model. The goal is to get a better understanding of the processes ongoing in the real world problem. Empirical descriptive research typically follow a cycle of conceptualization, modeling and validation. The empirical research is more concerned with finding a fit between the model and the real world problem. Empirical normative research is focused on finding a strategy or policy to improve the current situation or state. This branch is very small in scientific literature and it is problematic to produce strong verification procedures for claims. In real world problems it is nearly impossible to control all variables. Therefore, it is hard to connect a change in performance to controlling certain variables in a process. This type of research follows the complete cycle of conceptualization, modeling, model solving and implementation. It is common to build research on already existing theories from the axiomatic descriptive research. Our research project also follows the whole cycle like in the empirical descriptive research type. In addition, we also have an orientation phase as described by van Aken et al. (2012). The orientation phase is the first phase in our research project before we enter the cycle of Mitroff et al. (1974).

Orientation

The outcome of the orientation phase is the project proposal.

Conceptualization

In this phase we describe existing concepts in scientific literature that serve as a basis for our own model. The literature review provided an overview of different concepts used in supply chain network design (SCND). Chapter 4 contains a conceptual model that we chose for our SCND problem with mathematical formulation. Furthermore, it describes all relevant concepts that we use in our research project.

Modeling

The conceptual models from the review of scientific literature are the basis for our own model. We propose a quantitative model in mathematical terms to solve the problem at Alpha. We use a mixed integer linear programming (MILP) model to determine what locations are best suited for assembly of products under different future scenarios. Uncertainty about the future is included by using random variables in the MILP model formulation. This formulation is named stochastic programming which is a branch of mathematical optimization that includes uncertainty related to parameter values. Future scenarios are generated by using a common technique in stochastic programming, namely the sample average approximation technique.

Model solving

The model that we constructed needs to be solved for the specific values of parameters that reflect Alpha's situation. Solving the model is done by using a specific algorithm developed by researchers working on a SCND problem.

Implementation

The implementation phase chapter describes how the results should be used by Alpha.

2.3 Deliverables

This project have the following deliverables:

- Conceptual model: This model should be added value to the scientific literature
- Mathematical model and solution strategy: We transform the conceptual model into a mathematical model. This can be solved to obtain solutions for the problem. We provide an algorithm to solve the mathematical model.
- Decision support tool: The mathematical model is the basis for a decision support tool. Alpha can use this decision support tool to do analysis with specific parameter values.

2.4 Project scope

- The analysis is limited to the locations of the existing factories in Alpha's supply chain. Building new plants is out of the scope of this project's analysis.
- The input parameters for our mathematical model are sourced from various ERP systems. We have extracted cost parameters for labor cost, machine variable cost, material cost, transportation cost and machine fixed overhead cost. In addition we also took historical demand, working capital levels and lead times from the ERP systems.

Chapter 3

Supply chain network design

This chapter describes the concept of supply chain network design. First, we describe the simple models that form the basis for SCND. Hereafter, we provide a selection of supply chain management (SCM) aspects important for SCND. Finally, we discuss the concepts uncertainty and robustness and their importance in SCND.

3.1 Facility location models

In our literature review we found that facility location models are the basis for the SCND models. We studied the review of Melo et al. (2009) and Klibi et al. (2010), they both also confirm that this assertion. Therefore, we start this chapter with a section that explains these facility location models. The SCND models are basically an extension of these facility location models by adding SCM aspects. The most simple location facility models are the discrete versions in which all facilities have the same role. The goal is to minimize the total cost associated with serving a set of customers from the chosen facility locations. Setup costs for opening a new facility are equal for all locations. This is the so called p-median problem and it has been widely studied in the literature (Revelle et al., 2008). We have a set $\mathcal{I} = \{1, .., i, ..m\}$ with candidate locations and a set $\mathcal{J} = \{1, .., j, ..n\}$ of customer locations. Let d_j and c_{ij} denote respectively the demand at location j and the unit cost of satisfying customer j from facility i. The total number of facilities to locate is equal to p, x_{ij} is a binary variable indicating if demand for node j is satisfied by facility i. Finally, a binary variable y_i is introduced to indicate if facility i is opened. The complete mathematical formulation for the p-median problem is in equations (3.1)-(3.6).

$$\min \quad \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} d_j c_{ij} x_{ij} \tag{3.1}$$

$$\text{s.t.} \quad \sum_{i\in\mathcal{I}} x_{ij} = 1 \quad \forall j \in \mathcal{J} \tag{3.2}$$

$$\sum_{i\in\mathcal{I}} y_i = p \tag{3.3}$$

$$x_{ij} - y_i \leq 0 \quad \forall i \in \mathcal{I}, \quad \forall j \in \mathcal{J} \tag{3.4}$$

$$y_i \in \{0, 1\} \quad \forall i \in \mathcal{J} \tag{3.5}$$

$$x_{ij} \in \{0, 1\} \quad \forall i \in \mathcal{I}, \quad \forall j \in \mathcal{J} \tag{3.6}$$

When the setup costs are not equal for all locations, then the problem changes to a new type. This type is called uncapacitated facility location problem (UFLP). The above mathematical formulation can be easily adjusted for UFLP. If we change the objective function in equation(3.1) to equation (3.7), then we have the formulation for the UFLP. In equation (3.7) f_i denotes the setup cost for location i.

$$\min \quad \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} d_j c_{ij} x_{ij} + \sum_{i\in\mathcal{I}} f_i y_i \tag{3.7}$$

For both the p-median and UFLP, each customer is served from the closest facility. Another extension of the UFLP is the capacitated facility location problem (CFLP). This extension assigns maximum capacities to the facilities.

The previously described models are very simple in terms of their characteristics. They consider a single-period planning horizon, single product type, single type of facility, deterministic parameters and location-allocation decisions. These simple types are also studied most in Melo et al. (2009). They can be extended to create more complex models which are better suited to model the real world. Nonetheless, all of the more comprehensive models which have been developed are based on these simple model structures. When designing a supply chain it is important to use these more comprehensive models and add some more relevant aspects to the objective function.

3.2 Supply chain management aspects

Melo et al. (2009) analyze what type of SCM aspects have been included in the facility location models. The aspects related to inventory and production are the most included aspects in Melo et al. (2009). These aspects include determining stock locations, defining levels of inventory and setting production capacities in the network. Their analysis also includes networks with reverse logistics. Reverse logistics are the operations related to the reuse of materials. Businesses collect the products back from their customers to reuse it in their manufacturing process. The e-commerce industry introduced some specific reverse logistics to the supply chain management industry. In this industry businesses have a significant flow of product returns from their customers. Melo et al. (2009) argue that networks with reverse logistics are becoming more important, because of the growing e-business and environmental awareness which stimulates reuse of materials. A full closed-loop network which includes the forward and backward logistics is significantly more complex.

Globalization leads to more challenging decisions concerning supply chain network design. Financial aspects such as different tax rates and government incentives also have to be taken into consideration. We also take into account the tax costs for the different factories in our research project. Alpha's factories are in different countries and therefore subject to different tax rates. These aspects require for more robust and reliable network structures in order to stay competitive. The term robustness means in this case the ability of the supply chain network to perform well in different future scenarios. Supply chain's reliability refers to the capability of delivering constantly quality performance. For example, delivery reliability is a very important aspect in SCM.

Within SCM there exists three levels of planning horizons, namely: long-term, mid-term and short-term (Stadtler & Kilger, 2008). The long-term horizon decisions determine the structure of the supply chain network. The prerequisites for the network structure are defined and these decisions have long term impact on the organization. The mid-term planning horizon are typically ranges from 6 to 24 months. On this horizon decisions are made on rough production quantities and resources for facility locations. The short-term horizon decisions have the highest degree of detail. Examples are daily and weekly schedules for machines in factories or planning of working shifts. The long-term decisions often involve investing large amounts of money and shall have great impact on the business performance. This description shows there is a clear link between facility location models and SCM, especially on the long-term planning horizon.

3.3 Uncertainty

The deterministic model outcomes have no guarantee for future performance. This highlights the importance of including uncertainty into the modeling approach. Supply chain environments change and this means we need to account for this in the model. Klibi et al. (2010) make a distinction between three concepts. They characterize the decision-making situations by the quality of information that is available.

- Uncertainty
- Certainty
- Risk

When there is perfect information available, decisions can be made with certainty. On the opposite side there is a situation in which there is imperfect information and a decision has to be made with uncertainty. Stewart (2005) explains that uncertainty leads to risk and this refers to the possibility of undesirable outcomes. Klibi et al. (2010) describes three types of uncertainty:

- Randomness
- Hazards
- Deep uncertainty

The first one, randomness, is defined by events related to supply chain operations that lead to variation. Hazards are low probability high-impact events that partially or completely disrupt the supply chain. Deep uncertainty refers to events on which there is very little information available to assess the likelihood of happening. Klibi et al. (2010) state that there exists no model in the literature which fully covers all of these three sources of uncertainty in the process of SCND.

The concept of randomness can be found in some parameters like customers demand, exchange rates and raw material prices. There are various approaches to include these random parameters in the model. One of the most widely adopted approaches is stochastic programming with recourse (Birge & Louveaux, 2011). Recourse programs are those in which some decision or recourse has to be taken after disclosure of uncertainty. A two-stage stochastic programming problem is an example of this approach. For example, the first stage in SCND is a strategic decision on which facilities to open. The second stage or recourse decision is the operational decision on how to route each product through the supply chain network. Santoso et al. (2005) argue that most of the static deterministic models can be reformulated in terms of stochastic programming with recourse. This approach also allows for the modeling of risk aversion of the decision maker (Shapiro, 2007). A further explanation and investigation on stochastic programming is in Subsection 4.6.

Modeling hazardous events seems to be quite challenging and it has also not received much attention in the area of SCND. It should not be completely neglected, since the impact on the business can be catastrophic. The importance in SCND depends of course on the risk appetite of the decision maker. Several approaches for modeling hazards are proposed by Klibi et al. (2010), but none of them are superior compared to each other. This area of research needs some more attention in future research on SCND.

The inclusion of deep uncertainty in SCND can be done through qualitative methods as well as quantitative methods. Klibi et al. (2010) name the Delphi method as one of the methods used in practice by a company like Shell (Royal Dutch Shell Group, 2005). A quantitative method that is found in literature is robust optimization. Some robustness criteria that can be used are the minimization of maximum costs and the minimization of the maximum regret (Shapiro, 2007).

3.4 Robustness

Klibi et al. (2010) define three types of robustness in their review, namely: model robustness, algorithm robustness and decision or solution robustness. In the process of SCND we are definitely interested in the solution robustness. Rosenhead et al. (1972) state that robustness is: "a measure of the flexibility which an initial decision of a plan maintains for achieving near-optimal states in conditions of uncertainty". Robustness should be measured on the supply chain's ability to create sustainable value for its shareholders over a range of plausible future scenarios. An appropriate performance measure for this is the DRCF generated by the SCND. In addition to the DRCF, we also consider the consequences for the lead times. For the situation at Alpha, we will also search for possible future scenarios. Within these scenarios we want to determine the best configuration of the design for the supply chain, i.e. which configuration yields the most DRCF and best lead times for the possible scenarios. Stochastic programming uses this approach through the modeling of recourses. Risk aversion can be included by using a risk measure such as mean-variance instead of expected value. There also exist situations in which there is mixed information available about scenarios, so probabilistic and non-probabilistic. Robust optimization is an alternative approach in case there are no scenario probabilities available, i.e. deep uncertainty (Kouvelis & Yu, 1997). The fuzzy programming approach can be used to model uncertainty in a different way compared to stochastic programming.

The robustness of the design outcome is also related to two other important concepts in SCM: responsiveness and resilience. Responsiveness describes the ability of the supply chain to adapt to variations in the business-as-usual environment, such as changing demand. Examples of responsiveness enhancing policies are capacity buffers, safety stocks and subcontracting. In our research project at Alpha we consider the lead time performance for the SCND which is also directly related to responsiveness. Imagine a situation where customer places a big order for a product with some significant workload. This order would normally take longer than the standard lead time due to its size. If Alpha's manufacturing lead time is short and flexible, then they will be able to fulfill the wish of the customer without extending the delivery date. This shows how we aim to include responsiveness in our objective for the SCND. Resilience is the ability of the supply chain to quickly recover from a disruption or avoid them. For developing resilience it is the challenge to provide adequate protection against disruptions, but still remain effective in normal circumstances. The resilience of a supply chain can be done through the strategic design, such as choice of locations for production and inventory. The concept responsiveness can be based on flexibility or redundancy. Flexibility based capabilities can be obtained through setting up facilities which can execute multiple tasks within the supply chain. For example, setting up production facilities that can handle multiple type of products and operations. Redundancy based capabilities are the duplication of network resources in the design. An example would be the assignment of overcapacity throughout the final design of the network. In the literature there is a lack of stochastic models that include both responsiveness and resilience into their modeling approach (Klibi et al., 2010).

The robustness of a SCND is reflected by the risk mitigation constructs that are included. Low impact short-term risk mitigation refers to the concept of responsiveness. Resilience is the area of risk mitigation related to the structure of the network, i.e. the ability to cope with high-impact events.

Chapter 4

Conceptual model

The conceptual model serves as a basis for the mathematical model for SCND. In this chapter we set the scope of the model and identify the relevant parameters and variables.

4.1 Finance and supply chain

Figure 4.1 depicts an important relationship between finance and supply chain management. The firm uses funds to generate returns for its investors which are split into two groups, i.e. the debt and equity investors. In Figure 4.1 the cash flows from the bottom, where the investors are, up to the firm. From the firm the cash flows to operational assets and comes back to the firm. Debt investors provide the company with cash on agreed terms about interest rate and duration of the loan. The firm pays the loaners in cash for the interest and repayment of the debt. Equity investors pay the firm in cash in trade for a part of the ownership of the particular firm. The firm pays the equity investors in cash in the form of dividend or share buybacks. Through these investments a firm acquires the required funds, which are allocated by the firm to generate return from supply chain operations. Allocation of funds is done through investing in operational assets such as machines, raw materials and inventory. The operational assets are used in the conduct of the ongoing operations of the firm, this means in the case of Alpha the supply chain operations to serve its customers.

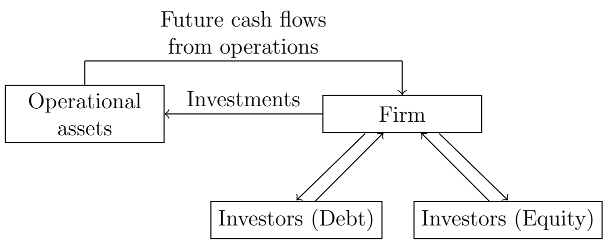

Figure 4.1: Cash flows in supply chain

Cash flows within the whole supply chain are depicted in Figure 4.2 with the firm in the center. This sketch is based on the description of Brealey et al. (2012). The customers pay the firm in cash for its delivered services and products. The operational expenses are paid in cash to contract manufacturers and third party logistics providers. In addition, the firm also pays operational expenses to for example its workforce, machine cost and overhead cost. The firm pays cash to the third party suppliers for raw materials and components required for operations.

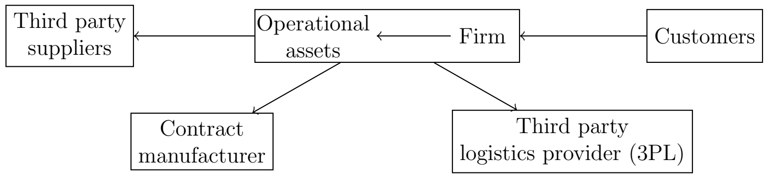

Figure 4.2: Cash flows in supply chain

We have to identify which cash flows are relevant for our own analysis, this means which cash flows are influenced by the SCND from our model. Changing the assembly locations influences the operating expenses, transportation cost and tax cost. We have specified the operating expenses for assembly operations of Alpha in Table 4.1. If we decided to change an assembly location, then we move the existing assembly facilities to a new location. When we choose to move the existing facilities to a new location it will mainly influence the labor cost. We assume that machine variable cost and fixed overhead remain the same independent of the location. Aside from the labor cost, the material cost is also influenced by the location of assembly, because in the current situation a part of the components are sourced from Alpha's own factories. If Alpha buys components from its own factories, then the component producing factory has to charge a minimum profit margin by law and its charged with profit tax. Finally, the transportation costs are dependent on the location of assembly, because components have to be transported to the location of assembly. We can avoid paying profit tax and transportation cost if we assemble the product on the same location as where the components are produced.

The most important cash flows which are influenced by our decision on the SCND are going from the firm to the third party logistics provider and the operating expenses. We assume for our model that the cash flow to contract manufacturers, to third party suppliers and from customers are not influenced by our decision. Within the firm in

Table 4.1: Operating expenses

Material cost	
Labor cost	
Machine variable cost	
Machine fixed overhead	+
Operating expenses	

Figures 4.1 and 4.2 there are also cash flows in our case. We have put a schematic drawing of the cash flows in the Alpha group in Figure 4.3. All entities on the bottom belong to the Alpha group and their sole purpose is to serve in the best interest of the Alpha group. The consequence of this is that Alpha would normally seek to remit earnings back to headquarters in the UK to service its shareholders. For investments in operational assets Alpha group transfers cash to the relevant factory. Alpha has a policy in place for the prices that Alpha companies charge each other, because they want to optimize the total return from the whole supply chain. Component and assembly factories charge a minimum profit margin for their products to follow this policy. In our model we assume for simplicity a fixed profit margin for all products. The relevant cash flows in Figure 4.3

are from the assembly factory to the components factory, from the components factory to Alpha group and from the assembly factory to Alpha group. They reflect respectively the situations where components are bought for assembly, profit from selling components is sent to Alpha group and profit from selling finished goods is sent to Alpha group. We assume that the other cash flows are not affected by the choice for a new SCND. Our model includes therefore only the components and assembly factories. We consider the hubs as the customers in our model. The hubs are the finished goods warehouses in Alpha's supply chain. From the hubs the products are sent to the end customer for a profit margin that is different depending on the country where the product is sold. We assume that the cost of components from third party suppliers are equal for each assembly location, which means that the cash flow from Alpha to third party suppliers remains unchanged.

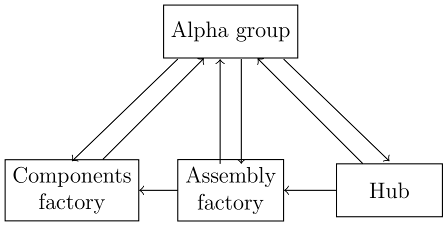

Figure 4.3: Schematic cash flows in Alpha group

4.2 Discounted residual cash flow

In our main research question we state that Alpha wants to create more value from their supply chain operations. The concept of value sounds rather vague and we will give it a specific definition before continuing. Value is in this case equal to the discounted residual cash flow (DRCF) from supply chain operations and the lead time performance. Below this paragraph is a formulation of the calculation of the DRCF. It starts with the revenue from sales minus the operating expenses. In our case we have direct labor, machine variable cost, machine fixed overhead and raw material as operating expenses. Now, we can execute the first calculation of the DRCF model below to arrive at the earnings before interest and taxes (EBIT). All of Alpha's factories pay profit tax when they sell a product to another entity, it doesn't matter if this is the customer or a supply chain factory of its own division. Tax costs are deducted from EBIT to obtain the NOPAT figure. From NOPAT we deduct changes in working capital (decrease) to obtain the RCF from operations. The final step is discounting the RCF at the weighted average cost of capital (WACC). The result is the DRCF or in other words a part of the objective function for our model. We use a two year horizon for our analysis. This means we have to do the calculation below for two years with the appropriate numbers.

Revenue - Operating expenses = EBIT (Earnings before interest and taxes)

EBIT - Taxes on EBIT = NOPAT (Net operating profit after taxes)

NOPAT - Changes in working capital

= RCF (Residual cash flow)

$$DRCF = \frac{RCF}{1 + WACC}$$

The above formula captures all performance measures that Alpha defined explicitly except the lead time performance. We believe that a shorter lead time can give higher revenues, because it makes the company more competitive. Therefore, we add the lead time performance to the DRCF in a weighted objective function with varying weights. In Section 4.5 we will further elaborate on the lead time performance.

Alpha is a publicly traded company and its paramount goal is to maximize its shareholders wealth. The weighted objective function is in our opinion the best way to create value from the new SCND for its shareholders. An increase in residual cash flow enables the firm to return cash to its shareholders through dividends or share buybacks. Better lead time performance can increase the company's competitive strength. The subsequent subsection provides information about the WACC and the calculation for this particular case.

4.3 Weighted average cost of capital

In corporate financing the source of financing is either equity, debt or both. Using debt to finance business activities is a straightforward approach, the business borrows money on agreed terms for interest rate and payback term. For this type of financing the credit investor bears not all the risk, because if the business fails, then he can recover fully or partially his investment in the form of the loan. Equity investing is an alternative form of financing in which investors provide the funds for financing in exchange for some ownership in the business. Equity financing is used by all publicly traded companies and investors are the buyers of the offered shares at the stock exchanges. The investors are bearing the full risk in this type of financing, because if the company goes bankrupt they most probably end up with nothing. Most companies use a combination of debt and equity financing, they borrow money and sell ownership through stocks of some form.

There are arguments in favor for both forms of financing, but that is beyond the scope of this research project. We accept it is a given fact that Alpha's financing structure is a combination of debt and equity financing. Both types of financing have different associated costs. The calculation of the cost of debt takes into account the interest rate and tax rate, because interest expenses are tax-deductible. The formula for the cost of debt:

$$\text{Cost of debt} = C_d = \text{Interest expenses} \times (1 - \text{Tax rate}) \tag{4.1}$$

The cost of equity is somewhat more complicated to calculate, because there is no technical price like the interest rate for cost of debt. Investors have a certain expected rate of return on their invested money, but it is not easy to calculate this in a straightforward manner. We use the capital asset pricing model (CAPM) to determine the cost of equity for Alpha in this case. Despite failing empirical experiments and some strong counter arguments (Fama & French, 2004), we think the CAPM is our best option to come up with an estimate for the cost equity. The CAPM enables us to do scenario analysis for the cost of equity.

$$\text{WACC} = \frac{V_e}{V_e + V_d} C_e + \frac{V_d}{V_e + V_d} C_d (1 - T_r) \tag{4.2}$$

We defined the formula for WACC in equation (4.2), see Table 4.2 for the definition of the symbols. All information required except for the cost equity we can obtain from the

Table 4.2: WACC formula symbols

Symbol	Definition
C_d	Cost of debt
C_e	Cost of equity
V_e	Value of equity
V_d	Value of debt
T_r	Tax rate
$V_e + V_d$	Market value of the firms financing

annual report of Alpha. In the subsequent subsection we will explain more about the CAPM, which we use to obtain an estimate for the cost of equity.

Capital asset pricing model

The CAPM assumes there are two types of risks associated with investing in an asset, namely the diversifiable and non-diversifiable risk. The first refers to risk that can be eliminated when spreading your investments across multiple types of assets, i.e. having an investment portfolio. The latter refers to the risk that is inherent to investing your money in any kind of asset, it is also known as the systematic or market risk. In addition to these two types of risk, the CAPM assumes there is such a thing as a risk-free investment that gives a guaranteed return. It is doubtful to assume that there exist such investments, but a US treasury bond is considered to be one of the safest investments one can make on financial markets. The last important element we need to explain is the β associated with a certain asset. This β reflects the assets sensitivity to the fluctuation in overall market changes, so high β means the asset responds sensitive to changes in the overall market. Low β means the asset is insensitive to changes in the overall market. Another assumption associated with the CAPM is that prices are set in an informationally efficient market.

Below in equation (4.3) is the formula to determine the expected rate of return for an underlying asset according the CAPM. The formula consists of two main components. We define the two components as the time-value of money and risk-premium for investing in a particular asset, further explanation on these components is below the formula.

$$\mathbb{E}[R_\alpha] = R_{risk-free} + \beta_\alpha(\mathbb{E}[R_{market}] - R_{risk-free}) \tag{4.3}$$

$$\beta_\alpha = \frac{Cov(R_\alpha, R_{market})}{Var(R_{market})} \tag{4.4}$$

In equation (4.3) $\mathbb{E}[R_\alpha]$ denotes the expected rate of return for asset α. The risk-free rate in the formula is denoted by $R_{risk-free}$ and is the time-value of money associated with the concept of net present value. An investor is certain to receive this return on his investment in the market and it has the lowest risk associated with it. The risk-premium component in the formula is defined as $\beta_\alpha(R_{market} - R_{risk-free})$, which expresses the additional risk compensation for this particular asset. This risk-premium is straightforward to explain, because we deduct the risk-free rate from the expected market return, which is the ordinary risk-premium for investing in the market. Finally, we multiply this ordinary risk-premium by the β_α coefficient to adjust it for the underlying asset's sensitivity for changes in the market's return.

In the CAPM β_α is defined as the covariance of the return from asset α and return from the market divided by the variance of the return from the market. We use linear regression with the ordinary least squares (OLS) method to determine the value of β_α. This approach supplies us with some statistical measures to judge the quality of the estimate. In the linear regression results we are interested in the value of the slope, which also happens to be denoted by the symbol β. Furthermore, we are interested in the t-statistic and confidence interval of the β from the linear regression. Aside from these statistics, we use the R-squared statistic, because this shows how much of the the asset's movements can be explained by the market's movements. In Section 6.5 we elaborate further on what data we are using for the CAPM.

4.4 Working capital

The SCND influences the lead time for the product families and this also has consequences for the amount of working capital that is required. We refer to the working capital that is in work in progress (WIP). For longer lead times we will need higher levels of WIP to produce at the same rate. Our calculation for required WIP is based on Little's law, which is a well known formula in queuing theory, see equation (4.5).

$$W = A_r L \tag{4.5}$$

The symbols W, L and A_r denote respectively the average number of jobs, the average sojourn time and the rate at which jobs arrive at the system. In our analysis the average number of jobs in the system is equal to WIP and the average sojourn time is the lead time. The rate at which jobs arrive at the system does not change when we change our SCND, so we will simply treat this as a given number. We can influence the sojourn time which is the lead time in our model.

We use Little's law to calculate the difference in WIP for different assembly locations scenarios. In equations (4.6)-(4.8) is an example of how to calculate the difference in WIP for different assembly locations. Let W_1 and L_1 denote WIP level and lead time for option 1, which is in this example the assembly location in the current situation. The symbols W_2 and L_2 denote WIP level and lead time for option 2 and this is the alternative assembly location in the example. The difference in WIP is straightforward to calculate by using equation (4.8). Finally, we multiply the difference by the percentage of components that come from the alternative location. For example, when currently 70% of the components are sourced from the location of option 2, then we multiply $W_{difference}$ by 0.7. We use this result as "Changes in working capital" in the DRCF, i.e. in the objective function. Let us remark that a positive number for $W_{difference}$ means that the working capital decreases.

$$W_1 = A_r L_1 \tag{4.6}$$

$$W_2 = A_r L_2 \tag{4.7}$$

$$W_{difference} = W_1 - W_2 = A_r(L_1 - L_2) \tag{4.8}$$

The calculation in equation (4.5) is valid for the long run average state of a system. We use Little's law to estimate the effects on the working capital in our different configurations of the supply chain.

4.5 Lead time performance

The optimization problem for this SCND project also includes the lead time performance difference for the various future scenarios. Let us describe what performance measure we use for the lead time performance. We argue that it is best to look at the weighted relative performance difference compared to the current situation. For example, if we can reduce lead time with 2 days and the total manufacturing lead time for a product is 4 days, then this means a reduction of 50%. In addition, we should weigh the importance of this reduction by including the value of this product family compared to others in Alpha's portfolio. So when we multiply the relative change by the total value of the product family we have our weighted lead time performance difference. Refer to Section 6.3 for the exact formula. We define the value of a product as the total manufacturing cost multiplied by the total demand of a product family for the current situation. The total demand is the sum of the demand in all time periods included in the analysis. This means that value can vary, because the total demand can be different. We combine this performance measure with the DRCF in our final optimization model.

For each product family we determined where components are sourced from. We aim to improve the lead time performance by eliminating the transportation lead time of components between factories. If components are sourced from only one other Alpha factory, then we can improve the lead time performance. In contrast, when components are sourced from multiple Alpha factories, we are not able to improve the lead time performance. Finally, we assume that parts sourced from third parties do not influence lead time performance. We give an example to describe what we mean by this definition. Consider a product family which sources components from Alpha's Jupiter factory and it is currently assembled in the factory on location Saturn. If we change the assembly location to location Jupiter, then we can shorten the lead time and thus improve lead time performance. The situation changes if the factory receives components from location Jupiter and Uranus. We can not improve the lead time performance if we move the assembly operations to either location Jupiter or Uranus. If we moved the assembly location to one of the sources of the components, then there is still some lead time for components coming from the Jupiter or Uranus factory. The improvement in lead time performance basically is the reduction in transportation lead time from the components machines to the assembly machines. Contrary to the improvement, we can also design a supply chain network with longer lead times. The situation where we lengthen the lead time is possible if the products are currently assembled on the same location as where all of its components are produced. The lead time would increase if we move the assembly operations to another location, because there is now transportation time for components involved. This sketches the trade off that is made in the designing process between DRCF and lead time performance.

4.6 Stochastic programming

The problem for this research project is suited for using stochastic programming techniques. Stochastic programming is an extension of linear programming, which is a common technique in the operations research field. The extension is the introduction of random variables in the linear programming formulation. In our project we are dealing with some unknown values for parameters in the future. To be able to make the best decision we want to analyze different scenarios. In the next subsection we introduce the basic principles of the stochastic programming technique.

Basic formulation

Consider the basic linear programming problem in (4.9)-(4.11). This problem can be solved straightforward with standard techniques, such as the simplex algorithm.

$$\min \quad z = c^T x \tag{4.9}$$
$$\text{s.t.} \quad Ax = b \tag{4.10}$$
$$x \geq 0 \tag{4.11}$$

Birge & Louveaux (2011) describes stochastic linear programs as linear programs in which some of the included variables are uncertain. Recourse programs are those in which some decision or recourse has to be taken after disclosure of uncertainty. From here we assume that the random variable follows a known probability distribution. The values of the random variables are only known after the random experiment. This means that the realization of the vector $\boldsymbol{\xi} = (\boldsymbol{q}, \boldsymbol{T}, \boldsymbol{W}, \boldsymbol{h})$ is only known after the experiment. The symbols $\boldsymbol{q}$, $\boldsymbol{T}$, $\boldsymbol{W}$ and $\boldsymbol{h}$ represent the random variables with known probability distributions. In our formulation $\boldsymbol{\xi} = (\boldsymbol{q}, \boldsymbol{T}, \boldsymbol{W}, \boldsymbol{h})$ denotes the random data vector, while $\xi = (q, T, W, h)$ represent a particular realization (scenario) of this random data vector. The set of decisions is split into two groups:

- First-stage decisions: These decisions have to be taken before a realization of the random variables. They are taken in the first-stage.
- Second-stage decisions: After the realization of the random variables this group of decisions is made. These decision are taken in the second-stage.

The classical formulation of stochastic programming with fixed recourse is given in (4.12)-(4.14).

$$\min_{x \in S} \quad f(x) = c^T x + \mathbb{E}[Q(x, \boldsymbol{\xi})] \tag{4.12}$$
$$\text{s.t.} \quad Ax = b \tag{4.13}$$
$$x \geq 0 \tag{4.14}$$

where $Q(x, \xi)$ is the optimal value of the second-stage problem in (4.15)-(4.17).

$$\min_{y \in V} \quad q(\xi)^T y \tag{4.15}$$
$$\text{s.t.} \quad T(\xi)x + W(\xi)y = h(\xi) \tag{4.16}$$
$$y \geq 0 \tag{4.17}$$

For the above problem we have that $\xi(q, T, W, h)$ is the realization of the random variables in the second-stage. After the random variables are realized, the two-stage problem boils down to a standard linear programming problem. We can use standard optimization techniques to arrive at the optimal solution. The presented formulation is a basis and can be extended to create more complex problems. For example, we can add integrality constraints to get a mixed integer linear programming problem.
The complexity lies in the optimal decision for the first-stage. In the above formulation we would have to evaluate expected value of the objective function in (4.12) for a feasible

choice for x. This can lead to complex calculations depending on how much random variables are included and their probability distribution functions. In addition, when there are a lot of options for the decision variable this gives a big number of possibly complex calculations. In the literature related to stochastic programming there are certain techniques proposed to deal with this problem. We describe the sample average approximation algorithm to deal with this difficulty in the next subsection.

Sample average approximation

Kleywegt et al. (2002) describe the method of the sample average approximation (SAA). The method tries to estimate the optimal solution by taking samples of realizations of random variables. By solving the subproblems that arise from the samples we can estimate what would be the optimal solution for the stochastic programming problem.
Consider the optimization problem in (4.12). We can try to approximate this "true" problem by the problem:

$$\min_{x\in S} \left\{ \hat{f}_N(x) := c^T x + \frac{1}{N}\sum_{n=1}^{N} Q(x,\xi^n) \right\} \tag{4.18}$$

In (4.18) ξ^n is a vector in the generated random sample $(\xi^1, ..., \xi^N)$ with sample size N. The set of all feasible solutions is denoted by S. Let S^* and $\hat{S}_N$ denote respectively the sets of optimal solutions for problems (4.12) and (4.18). Let v^* and $\hat{v}_N$ denote the optimal values of the respective problems below.

$$v^* := \min_{x\in S} f(x), \quad \hat{v}_N := \min_{x\in S} \hat{f}_N(x)$$

Statistical inference

Kleywegt et al. (2002); Mak et al. (1999) and Norkin et al. (1998) give some useful properties for the SAA. These properties show it is an appropriate method for determining a "good" close to optimal solution for the "real" problem. In other words, the method is suited to obtain solutions that are close to optimal with a predefined optimality gap ϵ. In statements hereafter the abbreviation w.p.1 means with probability one. They proofed the properties enumerated below this paragraph. The first two are related to the convergence of the objective values and solutions. Property three is about the convergence rate of $\hat{v}_N$ and $\hat{S}^\delta_N$ to their true counterparts. Mak et al. (1999) and Norkin et al. (1998) provide a statistical lower bound for v^*, which is the optimal value of the "true" problem. We show this in the fourth property.

1. $\hat{v}_N \to v^*$ w.p.1 as $N \to \infty$
2. We consider the ϵ-optimal solutions. That means, for $\epsilon \geq 0$, we say that $\bar{x}$ is an ϵ-optimal solution if $f(\bar{x}) \leq v^* + \epsilon$. The sets of ϵ-optimal solutions of $\min_{x\in S} f(x)$ and $\min_{x\in S} \hat{f}_N(x)$ are denoted by S^ϵ and $\hat{S}^\epsilon_N$, respectively. For any $\epsilon \geq 0$ the event $\{\hat{S}^\epsilon_N \subset S^\epsilon\}$ happens w.p.1 as $N \to \infty$
3. If δ is a number such that $0 \leq \delta \leq \epsilon$, then $S^\delta \subset S^\epsilon$ and $\hat{S}^\delta_N \subset \hat{S}^\epsilon_N$. Let S^δ and $\hat{S}^\delta_N$ denote respectively the sets of δ-optimal solutions of $\min_{x\in S} f(x)$ and $\min_{x\in S} \hat{f}_N(x)$.

From proposition 2 it follows that the event $\{\hat{S}_N^\delta \subset S^\epsilon\}$ happens w.p. 1 for N large enough. Under mild regularity conditions and $\delta \in [0, \epsilon]$, the probability of event $\{\hat{S}_N^\delta \subset S^\epsilon\}$ approaches 1 exponentially fast as $N \to \infty$

4. For any subset S' of S inequality $\hat{v}_N \leq \min_{x \in S'} \hat{f_N}(x)$ must hold. When $S' = S^*$ we have $\hat{v}_N \leq \min_{x \in S^*} \hat{f_N}(x)$, hence $\mathbb{E}[\hat{v}_N] \leq \mathbb{E}\left\{\min_{x \in S^*} \hat{f_N}(x)\right\} \leq \min_{x \in S^*} \mathbb{E}[\hat{f_N}(x)] = v^*$, where S^* is the set of optimal solutions of $\min_{x \in S} f(x)$.

In Section 5.2 we introduce the SAA algorithm which Santoso et al. (2005) use in their SCND problem. We are going to apply it to our own SCND, which differs slightly from their formulation, but the algorithm is still applicable.

Multi-objective linear programming

We defined two objective functions, i.e. the DRCF and the weighted relative lead time performance difference. There is a slight adjustment needed for equation (4.9) to include both objectives. The reformulation is:

$$\min \quad z = (c_1^T \alpha_1 + c_2^T \alpha_2)x \tag{4.19}$$

The new formulation in (4.19) is a bi-objective linear programming problem. We can normalize the weights $\alpha_1 + \alpha_2 = 1$ without loss of generality, because there are only two objectives. This can be further simplified by introducing one new variable λ to replace both α_1 and α_2. The new objective function is defined in the range $0 \leq \lambda \leq 1$ as:

$$\min \quad z = (c_1^T \lambda + c_2^T (1 - \lambda))x \tag{4.20}$$

Chapter 5

Modeling

In this chapter we propose a mathematical model for optimizing the SCND. We first formulate the deterministic linear programming variant of our problem. Hereafter, we include random variables in this formulation and define the stochastic programming formulation. The last part describes the algorithm for solving our stochastic programming problem.

5.1 SCND formulation

We propose a model formulation that is very much related to that of Santoso et al. (2005). The linear programming formulation is based on a network with a set of nodes $\mathcal{N}$ and arcs $\mathcal{A}$. There are three types of nodes, namely customers, suppliers and manufacturers. We have $\mathcal{N} = \mathcal{S} \cup \mathcal{P} \cup \mathcal{C}$ which means that the set of nodes consists of suppliers, production locations and customers. Our production locations $\mathcal{P}$ are the assembly locations and the suppliers $\mathcal{S}$ are the components factories. The set $\mathcal{K}$ denotes the set of all products produced. Let set $\mathcal{T}$ denote the set containing all time periods included in our analysis. The decision variables are x_{ij}^{tk} and y_i^k. These respectively are the flow quantity of product k from node i to node j in period t and the binary variable indicating whether node i is used for manufacturing product k. The flow quantity means the amount of product type k that goes from node i to node j in time period t.

$$
\begin{aligned}
\min \quad & \lambda\left(\sum_{k\in\mathcal{K}}\sum_{i\in\mathcal{P}} l_i^k y_i^k\right) + (1-\lambda)\left(\sum_{k\in\mathcal{K}}\sum_{i\in\mathcal{P}} c_i^k y_i^k + \sum_{t\in\mathcal{T}}\sum_{k\in\mathcal{K}}\sum_{(ij)\in\mathcal{A}} q_{ij}^{tk} x_{ij}^{tk}\right) && (5.1)\\
\text{s.t.} \quad & y \in Y \subseteq \{0,1\}^{|\mathcal{P}|} && (5.2)\\
& \sum_{j\in\mathcal{C}} x_{ij}^{tk} \le x_{ni}^{tk} \quad \forall\, n \in \mathcal{S}, \quad \forall\, i \in \mathcal{P}, \quad \forall\, k \in \mathcal{K}, \quad \forall\, t \in \mathcal{T} && (5.3)\\
& \sum_{i\in\mathcal{P}} x_{ij}^{tk} \ge d_j^{tk} \quad \forall\, j \in \mathcal{C}, \quad \forall\, k \in \mathcal{K}, \quad \forall\, t \in \mathcal{T} && (5.4)\\
& \sum_{j\in\mathcal{C}} x_{ij}^{tk} \le M y_i^k \quad \forall\, i \in \mathcal{P}, \quad \forall\, k \in \mathcal{K}, \quad \forall\, t \in \mathcal{T} && (5.5)\\
& \sum_{i\in\mathcal{P}} y_i^k = 1 \quad \forall\, k \in \mathcal{K} && (5.6)\\
& \lambda \in [0,1] && (5.7)\\
& x \in \mathbb{R}_+^{|\mathcal{A}|\times|\mathcal{K}|\times|\mathcal{T}|} && (5.8)
\end{aligned}
$$

The objective function consists of three components, see equation (5.1). The first component is the relative weighted lead time performance difference associated with the choice for assembly location. We denote the relative weighted lead time performance difference for assembly location i for product k by l_i^k. The investment cost is the second component. Variables c_i^k and y_i^k represent respectively the investment required to operate from node i for product k and the binary variable indicating if node i is used for assembly of product k. Aside from lead time performance difference and investment cost, the objective function contains the per-unit processing cost denoted by q_{ij}^{tk}. This per-unit processing cost is multiplied by the number of products that are going from node i to node j in period t for product k. Constraint (5.2) ensures that the variable y_i^k will be a binary variable. Constraint (5.3) ensures that every supplier delivers enough parts to node i to manufacture the amount shipped to customers for product k in period t. The next constraint, (5.4), ensures that node i satisfies all demand for product k in period t for all customers. A production node can only produce goods if it is “open”, which means that the corresponding binary variable y_i^k is equal to one. For this reason constraint (5.5) is included in the model. We only want one manufacturing node to be opened for each product, this is accomplished by constraint (5.6). Constraint (5.7) handles the multi-objective weights, which should always be between one and zero. The last constraint (5.8) ensures that variable x_{ij}^{tk} only consists of real numbers. We summarized all symbols and their definitions in Table 5.1.

Table 5.1: Definition of symbols

Sets	
Symbol	*Definition*
$\mathcal{T}$	Set of time periods
$\mathcal{S}$	Set of suppliers
$\mathcal{P}$	Set of possible production/assembly locations
$\mathcal{C}$	Set of possible customer locations
Parameters	
Symbol	*Definition*
q_{ij}^{tk}	The per-unit cost of letting product k flow in time period t from node i to node j
l_i^k	The weighted relative lead time performance difference when assembling product k at location i
d_j^{tk}	Demand for product k in time period t at node j
M	Big number
Decision variables	
Symbol	*Definition*
y_i^k	Binary variable which indicates if node i is used for the assembly of product k
x_{ij}^{tk}	The amount of product type k which goes from node i to node j in time period t
λ	The objective weight for the weighted relative lead time performance difference

Final formulation

The model in (5.1)-(5.8) is a standard MILP model with deterministic parameter values. We reformulate this model to obtain the stochastic programming formulation, because we want to include uncertainty related to certain parameter values. In (5.9)-(5.10) is our two-stage stochastic programming formulation where we treat the demand and cost as random variables with known probability distribution. The vector $\boldsymbol{\xi} = (\boldsymbol{q}, \boldsymbol{d})$ represents the random data vector, a realization (scenario) of this vector is denoted by $\xi = (q, d)$. In the first stage we have to decide for every product which nodes are open for the assembly and what the objectives weights are. The decision variable for opening an assembly node at location i for product k is denoted by y_i^k. We set the objectives weights values by choosing a feasible value for λ in the first stage. For the second stage, we have to satisfy all demand by using the SCND that was chosen in the first stage. The second stage optimization problem is formulated in (5.11)-(5.16).

$$\min_{y} \quad \{f(y,\lambda) := (1-\lambda)\sum_{k\in\mathcal{K}}\sum_{i\in\mathcal{P}} c_i^k y_i^k + \mathbb{E}[Q(y,\lambda,\boldsymbol{\xi})]\} \tag{5.9}$$

$$\text{s.t.} \quad y \in Y \subseteq \{0,1\}^{|\mathcal{P}|} \tag{5.10}$$

where $Q(y, \lambda, \xi)$ is the optimal value of the following problem

$$\min_x \quad (1-\lambda)(\sum_{t\in\mathcal{T}}\sum_{k\in\mathcal{K}}\sum_{(ij)\in\mathcal{A}} q_{ij}^{tk} x_{ij}^{tk}) + \lambda(\sum_{k\in\mathcal{K}}\sum_{i\in\mathcal{P}} l_i^k y_i^k) \tag{5.11}$$

$$\text{s.t.} \quad \sum_{j\in\mathcal{C}} x_{ij}^{tk} \le x_{ni}^{tk} \quad \forall\, n\in\mathcal{S}, \quad \forall\, i\in\mathcal{P}, \quad \forall\, k\in\mathcal{K}, \quad \forall\, t\in\mathcal{T} \tag{5.12}$$

$$\sum_{i\in\mathcal{P}} x_{ij}^{tk} \ge d_j^{tk} \quad \forall\, j\in\mathcal{C}, \quad \forall\, k\in\mathcal{K}, \quad \forall\, t\in\mathcal{T} \tag{5.13}$$

$$\sum_{j\in\mathcal{C}} x_{ij}^{tk} \le M y_i^k \quad \forall\, i\in\mathcal{P}, \quad \forall\, k\in\mathcal{K}, \quad \forall\, t\in\mathcal{T} \tag{5.14}$$

$$\sum_{i\in\mathcal{P}} y_i^k = 1 \quad \forall\, k\in\mathcal{K} \tag{5.15}$$

$$x \in \mathbb{R}_+^{|\mathcal{A}|\times|\mathcal{K}|\times|\mathcal{T}|} \tag{5.16}$$

5.2 Algorithmic Strategy

The difficulty of the stochastic programming approach is to determine an optimal solution. For our problem we use the sample average approximation scheme (Kleywegt et al., 2002; Mak et al., 1999) and (Norkin et al., 1998). This technique allows for solving a set of deterministic linear programming problems to obtain a solution with a predefined optimality gap. We try to estimate the value of $\mathbb{E}[Q(y, \lambda, \boldsymbol{\xi})]$, for this we generate $(\xi^1, ..., \xi^N)$ which is a sample with N scenarios. These samples are used to calculate $N^{-1}\sum_{n=1}^{N} Q(y, \lambda, \xi^n)$ which is our estimation of $\mathbb{E}[Q(y, \lambda, \boldsymbol{\xi})]$. The "true" problem in (5.9)-(5.10) is approximated by (5.17).

$$\min_{y\in Y} \quad \{\hat{f_N}(y,\lambda) := (1-\lambda)\sum_{k\in\mathcal{K}}\sum_{i\in\mathcal{P}} c_i^k y_i^k + \frac{1}{N}\sum_{n=1}^{N} Q(y,\lambda,\xi^n)\} \tag{5.17}$$

Lower bound (5.18) for the sample size is based on an absolute optimality tolerance $\delta \geq 0$, with $\epsilon > \delta$ and probability $\alpha \in (0, 1)$. It should be noted that this is a conservative bound. For practical use, one can often obtain solutions with similar optimality gaps by using smaller samples (Santoso et al., 2005).

$$N \geq \frac{3\sigma^2_{\max}}{(\epsilon - \delta)^2} \log(\frac{|Y|}{\alpha}) \tag{5.18}$$

In above notation, $\sigma^2_{\max}$ denotes the maximal variance of certain function differences, see Kleywegt et al. (2002) for details of the estimate (5.18). We will use a more efficient approach which repetitively solves smaller samples. Santoso et al. (2005) introduced this as the SAA algorithm and the details are provided in the subsequent subsection.

SAA Algorithm

1. *Step 1.* Generate M independent samples each with size N. This means we get samples $(\xi^1_j, ..., \xi^N_j)$ for $j = 1, ..., M$. For each sample we solve the associated SAA problem:

$$\min_{y \in Y,\ \lambda \geq 0} \quad \{\hat{f_N}(y, \lambda) := (1 - \lambda) \sum_{k \in \mathcal{K}} \sum_{i \in \mathcal{P}} c^k_i y^k_i + \frac{1}{N} \sum_{n=1}^{N} Q(y, \lambda, \xi^n_j)\}$$

   Let v^j_N and $\hat{y}^j_N$, $j = 1, ..., M$, respectively be an optimal objective value and an optimal solution.

2. *Step 2.* Compute

$$\bar{v}_{N,M} := \frac{1}{M} \sum_{j=1}^{M} v^j_N \quad \text{and} \quad \sigma^2_{\bar{v}_{N,M}} := \frac{1}{(M-1)M} \sum_{j=1}^{M} (v^j_N - \bar{v}_{N,M})^2$$

   In Section 4.6 we described a statistical lower bound for the optimal value of the "true" problem. So we have $\mathbb{E}[\bar{v}_{N,M}] \leq v^*$ where v^* is the objective value of the "true" problem. This is our statistical lower bound for v^*. In addition, the variance of this estimator is given by $\sigma^2_{\bar{v}_{N,M}}$.

3. *Step 3.* We can estimate the true objective function by calculating $\tilde{f}_{N'}(\bar{y}, \lambda)$ with a feasible solution $\bar{y} \in Y$ and a feasible choice for λ. An easy choice for $\bar{y}$ is a computed solution from $\hat{y}^j_N$ from *Step 1.*

$$\tilde{f}_{N'}(\bar{y}, \lambda) := (1 - \lambda) \sum_{k \in \mathcal{K}} \sum_{i \in \mathcal{P}} c_i \bar{y_i}^k + \frac{1}{N'} \sum_{n=1}^{N'} Q(\bar{y}, \lambda, \xi^n)$$

   The sample $(\xi^1, ..., \xi^{N'})$ is generated independently of the previously used samples. We choose N' larger than N, because we want to estimate the value of the true objective function for $\bar{y}$. Note that $\tilde{f}_{N'}(\bar{y})$ is an unbiased estimator of $f(\bar{y}, \lambda)$. The inequality $f(\bar{y}, \lambda) \geq v^*$ must hold because $\bar{y}$ is a feasible solution. This provides us with an upper bound of the value of v^*. If the sample $(\xi^1, ..., \xi^{N'})$ is iid (independent and identically distributed), then the variance of $\tilde{f}_{N'}(\bar{y}, \lambda)$ can be estimated as:

$$\sigma^2_{N'}(\bar{y}, \lambda) := \frac{1}{(N'-1)N'} \sum_{n=1}^{N'} ((1 - \lambda) \sum_{k \in \mathcal{K}} \sum_{i \in \mathcal{P}} c_i y^k_i + Q(\bar{y}, \lambda, \xi^n) - \tilde{f}_{N'}(\bar{y}, \lambda))^2$$

4. *Step 4.* By using the estimates from *Step 2* and *Step 3* we can calculate the optimality gap as

$$\text{gap}_{N,M,N'}(\bar{y}, \lambda) := \tilde{f}_{N'}(\bar{y}, \lambda) - \bar{v}_{N,M}$$

The variance of the optimality gap can be estimated by calculating

$$\sigma^2_{\text{gap}} = \sigma^2_{N'}(\bar{y}, \lambda) + \sigma^2_{\bar{v}_{N,M}}$$

The choice for $\bar{y}$ is naturally the solution vector from $\hat{y}^j_N$ with the smallest value for $\tilde{f}_{N'}(\bar{y}, \lambda)$.

In practice, the SAA algorithm is repetitively executed until it gives acceptable results for the optimality gap statistics. We follow this procedure in our case study for which the results are further elaborated in Chapter 7.

Chapter 6

Case study

This chapter elaborates on the case study that was performed during this thesis project at the company Alpha. We start by describing the current situation and the characteristics of this business case. Next, we define more precisely how we calculate the objective function. To conclude this chapter, we describe the modeling of random variables in our model.

6.1 Introduction

The objective function from the MILP model in Chapter 5 contains three parts, namely the investment cost (c_i), lead time performance difference (l_i) and the supply chain cost (q_{ij}). Figure 6.1 is schematic drawing of a routing for a particular product family. The lead time for this product family is defined as the the total time from the components factory until the finished product enters the hub. Two types of costs are in Figure 6.1 which are denoted by q_{ij} and q_{jk} and respectively represent the cost related to the supplier node and assembly node. In the next section we will further specify how the costs are calculated.

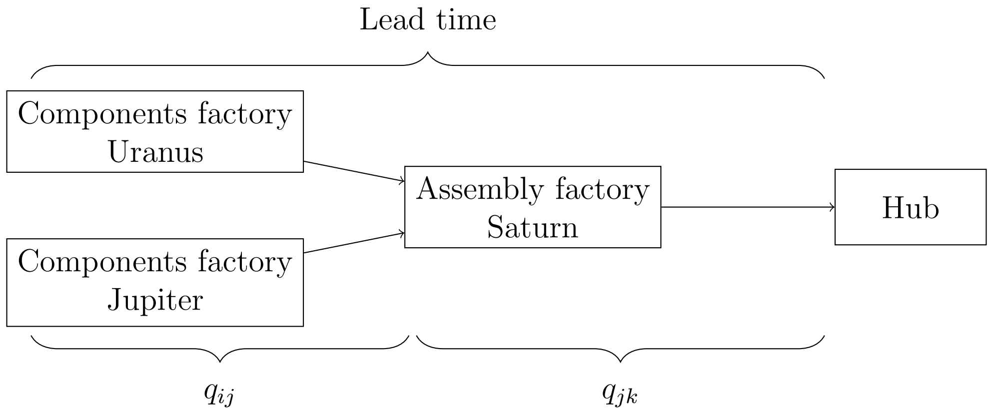

Figure 6.1: Schematic drawing of the production process (i, j and k respectively represent the components factory, assembly factory and hub)

6.2 Supply chain cost

The supply chain cost in our model consists of the cost related to manufacturing and transportation. For the assembly factory we have the following calculation:

Revenue - Operating expenses - Transportation expenses
= EBIT (Earnings before interest and taxes)
EBIT - Taxes on EBIT = NOPAT (Net operating profit after taxes)
NOPAT - Changes in working capital
= RCF (Residual cash flow)

The operating expenses are further elaborated in Table 6.1. In our case study we assume that the machine variable cost and machine fixed overhead cost remain the same independent of the chosen assembly location. The labor cost is different for each location and the material cost can be different and is dependent on the source of components for the product family. For the labor cost, we have gathered data related to the wages and labor cost for each country where Alpha has factories. We use this data to calculate conversion rates for each factory, which represent how much the labor cost would be on the new assembly location compared to the current assembly location.

Table 6.1: Operating expenses

Material cost	
Labor cost	
Machine variable cost	
Machine fixed overhead	+
Operating expenses	

The material cost can be different and this depends on the source of the components for the product family. For example we take the situation in Figure 6.1. In this situation there is transportation cost involved to deliver components to the assembly factory. If we change the assembly factory's location to one of the factories which currently produces components, then we can save this transportation cost. We assume that the receiving entities are paying for the transportation, so the assembly factory pays the transportation cost of components. Aside from transportation cost, we can also avoid paying profit tax for the components, because they are currently bought by the assembly factory from two different entities. If we assemble the products at one of the locations where components are produced, then we can eliminate paying profit tax for one of the components factories. The material cost for the assembly factory are also different, because currently the assembly factory is buying the components for a price which is higher than the cost price of the components. The components factory charges a profit margin which is set at a certain level and we assume this margin is fixed for all products for simplicity.

In the formula for the residual cash flow (RCF) the revenue is paid by the customers that buy the products, which are in our case the hubs and this also happens to be an Alpha entity. This implies that we do not strive for the highest possible revenue at the assembly nodes, because this drives the total cost for Alpha upwards. We are aiming to minimize the total cost for the Alpha's supplier factories, assembly factories and hubs.

The relevant cost for the hubs in our model is only the transportation cost, because this is influenced by our choice for the assembly location. We set our scope for this project to analyze the cost until it enters the hub. Our objective function for this case study is the total cost discounted at the WACC:

$$\begin{aligned}
&\text{Component factory cost} && = \begin{cases} \textit{Taxes on EBIT} \\ \textit{(NOPAT)} \end{cases} \\
&\text{Assembly factory cost} && = \begin{cases} \textit{Operating expenses} \\ \textit{Transportation expenses} \\ \textit{Taxes on EBIT} \end{cases} \\
&\text{Hub} && = \begin{cases} \underline{\textit{Transportation expenses} \quad +} \end{cases} \\
& && \quad\ \ \textit{Total cost}
\end{aligned}$$

In the objective function formulated above we take the sum of all relevant cost and discount it at the appropriate rate. The cost for the component factory is the tax on EBIT minus NOPAT from selling components. For the assembly factory, we explained what these cost represent in the beginning of this section. The transportation expenses for the hub are the cost paid for transporting finished goods to the hubs.

Transportation cost

In the former formulation of total cost we have transportation expenses for the components and finished goods. Alpha uses a third party logistics company for the transportation between factories and hubs. The first assumption is that Alpha's only hires complete trucks to do the transportation. In our model we assume the price per kilometer per truck is deterministic and is equal to £0.70. Furthermore, we assume that the truck's capacity is always 33 pallets. It would be too optimistic to assume that every truck is filled completely. Therefore, we assume that 90% of the truck's capacity is used in the long run for transportation. For a lot of products Alpha produces, we have collected data for how many of these products fit on one pallet. In addition, we also estimated how much components fit on one pallet. We combine all this information to calculate the transportation expenses for both the components and finished goods.

$$TC_i = D_i p_i^{-1}(33 * 0.9)^{-1} * c \tag{6.1}$$

The calculation for transportation cost (TC_i) for product family i in our model is in equation (6.1). The symbols D_i and p_i respectively represent the total demand for the next two years and the quantity per pallet for product family i. Finally, we have the cost per kilometer for a truck, which is denoted by c in (6.1).

6.3 Lead time performance

The lead time performance measure is denoted by l_i^k in the MILP in Chapter 5. We explained the concept of relative weighted lead time performance difference in Section 4.5. Now, we want to explain it by describing an example related to our case study. We are looking for potential savings on lead time through eliminating the transportation time of components. The example situation in Figure 6.1 is not suited to eliminate the transportation lead time, because there are two supplying factories for the components. If we adjust the situation in Figure 6.1 by removing for example components factory Uranus, then we can eliminate the components transportation time. This can be achieved when we move the assembly operations to location Jupiter, because now the components do not need to be transported to a different location for assembly. The final step is to multiply relative weighted performance difference by the current average supply chain cost times the demand. In equation (6.2) we define our calculation for the relative weighted lead time performance difference denoted by Y_i.

$$Y_i = (\frac{L_i^{transport}}{L_i^{manufacturing} + L_i^{transport}})\bar{V}_i D_i \tag{6.2}$$

The transportation lead time for components and current manufacturing lead time are denoted by respectively $L_i^{transport}$ and $L_i^{manufacturing}$. The first part of the formula for Y_i represents the relative lead time performance difference and this is multiplied by the defined weight. This weight is the current average cost price times the total demand in the upcoming two years, because our analysis has a horizon of two years. The demand in the next two years is denoted by D_i and $\bar{V}_i$ represents the current average cost price. This example shows how to calculate the relative weighted lead time performance difference for one product family.

6.4 Investment cost

A number of years ago Alpha decided to move some assembly operations to the Saturn location, because the labor cost were lower compared to the other locations. Over the years, the Saturn location improved the processes for the assembly operations and invested in more automation. Thereby they eliminated a part of the labor cost and increased their efficiency. If we decide to change a location for assembly operations, then we will move the existing facilities to a new location. We assume that there is no capital expenditure required to change the assembly location, but we take into account the new level of working capital, i.e. the WIP level. The investment cost (c_i) in our linear programming formulation is the difference in working capital. Furthermore, we assume that the difference in working capital is instantaneous and therefore we do not discount this at the WACC. In Section 4.4 we explained the method for calculating the difference in working capital for different assembly locations, see equations (6.3)-(6.5).

$$W_1 = A_r L_1 \tag{6.3}$$

$$W_2 = A_r L_2 \tag{6.4}$$

$$c_i = W_{difference} = W_1 - W_2 = A_r(L_1 - L_2) \tag{6.5}$$

6.5 Data

In this section we explain how we model the random variables in our model. We made a selection of product families for which enough useful data is available to model them in a proper way. After the selection we explain the modeling approach for the random variables.

Selection

We found that not for all product families there was enough data available to include them in our analysis. For some product families there was a very limited amount of data available for various reasons. There were also products for which the cost data was not appropriate to include them in this analysis. We have selected 79 out of the in total 187 product families, but these represent approximately 73% of value of all products sold in the last 12 months. This 73% is based on the cost price of the products not on the sales price.

We consider certain random variables related to demand, transportation cost and assembly cost. In the linear programming formulation from Chapter 5 we made a distinction between sets of products. These sets of products are Alpha's product families, which means there is a mix of products in the sets with different characteristics. We aggregate the historical data for the different product families and treat them in our model as random variables. Aggregating the data on product family level yields useful datasets as opposed to analyzing single items.

Operating and transportation expenses

There are differences between some of the products in the product families. For example, the labor and material cost can be different for products from the same family. In Figure 6.2 are three box plots for three elements of the assembly cost in a certain product family. Performing analysis on single item level is quite difficult, because there is a lot

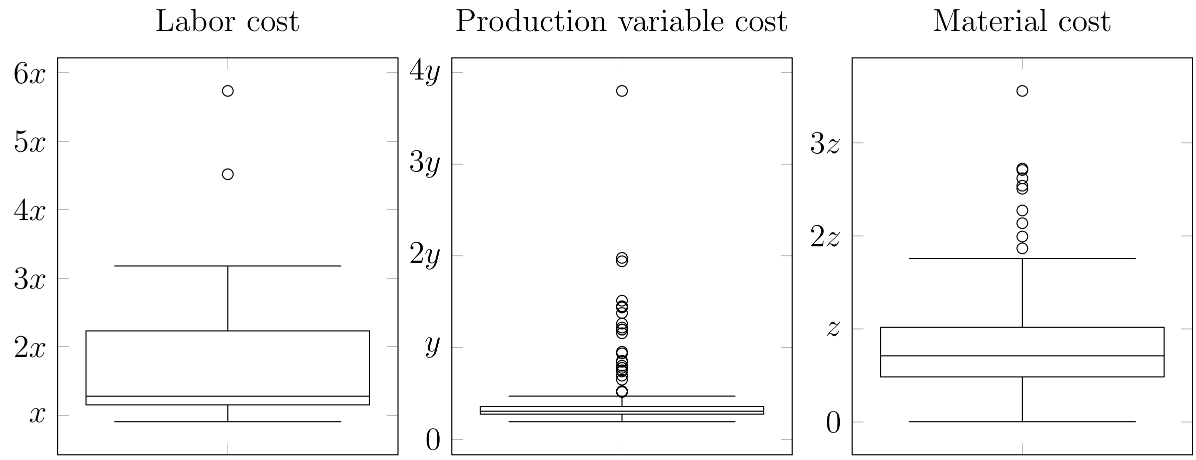

Figure 6.2: Box plots for assembly cost of product family

of variation on this level. Therefore, we choose to do our analysis on product family level. This means we look at historical data and calculate the average costs on product family level instead of doing this for every item. Figure 6.3 depicts the labor cost for the same product as in the box plots in Figure 6.2. It seems that most of the data on cost have a stationary probability distribution when aggregated on product family level. In addition, we assume that the random variables related to operating and transportation expenses are independent identically and distributed (iid). We want to model this data by fitting a theoretical probability distribution using a statistical package or use an empirical distribution when fitting gives poor results. We tried to fit theoretical probability distributions to the data that we collected. For the fitting we used the package *fitdistrplus*

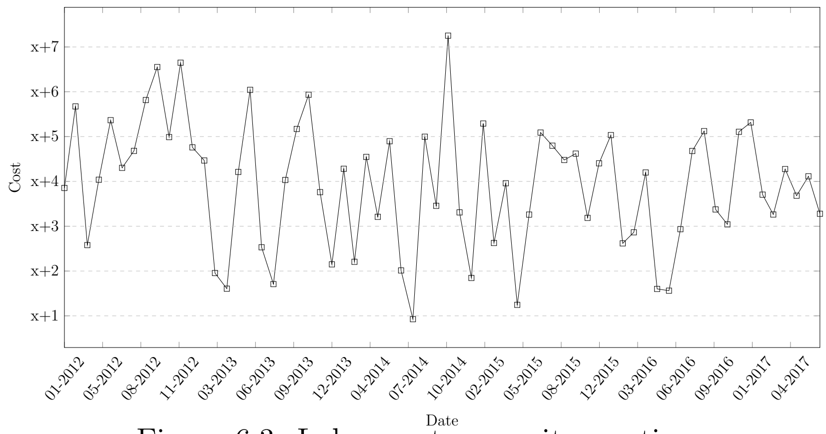

Figure 6.3: Labor cost per unit over time

in R. To demonstrate what we did we will give an example with labor cost depicted in Figure 6.3. We initialized the fitting procedure by looking at some descriptive statistics like skewness and kurtosis, see Figure A in the Appendix for a Cullen and Frey graph. This graph contains the statistics for the dataset itself and for a number of bootstrapped sets from the same dataset. From this graph we can get a rough insight what may be the appropriate theoretical distribution to fit on it. After this first inspection we run a fitting procedure which gave three goodness-of-fit statistics for a particular probability distribution. The estimation of the distribution parameters in *fitdistrplus* is done through maximum likelihood estimation (MLE). As a result from running this procedure we get goodness-of-fit statistics which are summarized in Table 6.2. We choose to try fitting the Log-normal, Gamma and Weibull distribution on our dataset. We exclude the normal distribution, because it can result in negative numbers which are not appropriate for cost data. The uniform distribution does not seem to fit our data set by visual inspection of a histogram. A Beta distribution supports only numbers between zero and one and this does by default not fit on this data set, because every data point is greater than one. From the goodness-of-fit statistics we conclude that the Log-normal and Gamma distribution are the best candidates. We will do some visual inspection of how well the theoretical distribution fits the data. In Figure 6.4 there is an example of the graphs we inspect to judge how well the fit is between the data and the Gamma distribution. In this case we can observe that the data fits good to the Gamma with the MLE parameters. The data set contains records for every month starting at January 2012. We are interested in annual cost data for our model and therefore we will draw twelve numbers from the distributions and use the mean value. If we collect the historical average annual cost, then we would have only five data points, which is not sufficient for a fitting procedure or too small for an empirical distribution. For the labor cost, we multiply the original

Table 6.2: Goodness of fit statistics

	Log-normal	*Gamma*	*Weibull*
Kolmogorov-Smirnov statistic	0.06449936	0.0636533	0.06812484
Cramer-von Mises statistic	0.05069189	0.0494266	0.05792592
Anderson-Darling statistic	0.31982556	0.3136035	0.55189854
Akaike's Information Criterion	-62.33656	-62.35664	-55.31214
Bayesian Information Criterion	-57.95725	-57.97733	-50.93283

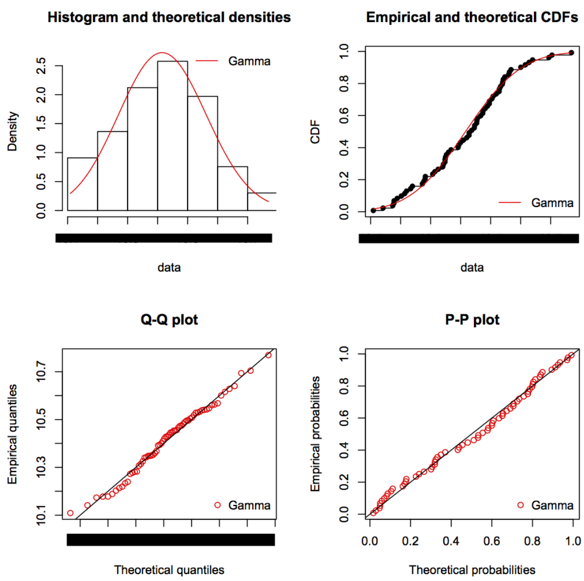

Figure 6.4: Goodness-of-fit graphs for Gamma distribution

number by a conversion rate, because the labor cost varies among the assembly locations' countries. The data that we are using for this calculation is sourced from Eurostat, which is an official bureau for statistics of the European Commission. We looked at the published data about the labor cost with the specific countries for Alpha's factories and used the median labor cost statistics.

Demand

The annual demand for a product family is modeled by the Log-normal distribution. We first tried to use an autoregressive integrated moving average (ARIMA) model to predict the future values, but this gave unrealistic outcomes. The prediction intervals associated with these models were so wide that we chose to not use these models. In this approach the problem was also caused by the limited size of the data set. The demand data set starts at January 2012 and contains monthly records. We used the monthly data for the ARIMA models, because the annual demand data would only contain five data points and this is not enough for a proper time series analysis. This means that we have to predict 24 periods ahead with monthly data points and this caused very wide prediction intervals. We use the expert knowledge of the demand planner at Alpha and chose a theoretical distribution with the appropriate parameters. The demand planner already had a two year forecast for all product families and a dataset with the forecast accuracy per product family for the past three years. With assistance of the demand planner we determined that the Log-normal distribution is a good distribution for modeling future demand. We split the two year forecast into two intervals of twelve months. For both these intervals we take the sum and set this as the expected demand for respectively the first and second upcoming year. The expected demand is also the expected value of the Log-normal distribution. The variance is specified according to the historical forecast error and expert knowledge about the volatility of the demand for certain product families. Now, we can sample numbers from the Log-normal distributions to model demand as a random variable. This is the total demand for these product families in the next two years, but we also want to specify how much demand each hub is consuming. We analyzed the demand specified per hub for the past five years and concluded that this is fairly stable over the long-run. For each product family we calculated the average percentage of demand that was going to each hub. This percentage is used to calculate the demand per hub for the next two years.

Cost of capital

For our analysis of the cost of capital we use the CAPM, the subsequent subsection provides more details on that model. The CAPM compares the return of a given asset against the return of the market. Our asset is listed on the FTSE 250, but we will use the S&P500 index as reference for the market. The S&P500 is an index that contains the 500 largest US companies based on their market capitalization. For the risk free-rate we use the yield from 10 year US treasury bills. It is open for discussion to use other references for the market data and risk free-rate. The historical data is the monthly percentage return of the asset and the market. We analyze the performance of Alpha's stock against the market by using historical data from between 24 and 36 months. We use OLS regression to determine the value of β from the CAPM and its confidence interval.

We started our analysis with 24 months of historical data and added one month at a time to historical data until we ended up with 36 months of historical data. Before adding another month to the historical data we performed an OLS regression analysis. The resulting model from each iteration is of the form:

$$y = \alpha + \beta x \tag{6.6}$$

The α and β in equation (6.6) are respectively named the intercept and the coefficient of the model. Monthly returns of the S&P500 hundred are denoted by x in (6.6). We are most interested in the value of β in (6.6) which also happens to be the estimate β for in the CAPM. From each OLS regression analysis we collected several statistics related to the coefficient, namely the estimated value, standard error, confidence interval and the p-value. The p-value statistic tells us with which probability we can assure that the coefficient is actually different from zero and thus has relevance in the model. Aside from the statistics for the coefficient we also collected the R-squared statistic which tells something about the complete model. The R-squared more precisely describes the proportion of variance in the dependent variable which can be explained by the independent variables, i.e. α and β in (6.6). We summarized the statistics in Table 6.3 from the OLS

Table 6.3: OLS analysis statistics

	Mean	*Standard error*	*Minimum*	*Maximum*
Model:				
R-squared	0.342431	0.020572	0.312177	0.387280
Coefficient:				
Estimated value	1.264311	0.057921	1.183791	1.371511
Standard error	0.336688	0.032694	0.301413	0.409017
p-value	0.001124	0.001056	0.000143	0.003841
Confidence interval upper bound	1.955446	0.111004	1.814255	2.170570
Confidence interval lower bound	0.573176	0.067728	0.471679	0.710185

regressions analyses. We observe that β is varying by some small margins and that the p-values all yield acceptable results. Our approach is to take the mean and standard error of the estimated values for β from the OLS analyses and model this a random variable which follows a normal distribution. This allows us to sample various values of β for the CAPM model and finally calculate different values for the WACC in our final model.

CAPM

We found an average return of the S&P500 of 9.33% measured from January 1988 until September 2017. The current yield on a 10 year US treasury bill is 2.21%. So we set the values of $\mathbb{E}[R_{market}]$ and $R_{risk-free}$ to respectively 9.33% and 2.21%. We model β_α from equation (6.7) as a random variable following a normal distribution with parameters $\mu = 1.264311$ and $\sigma = 0.336688$ based on statistics in Table 6.3. The mean value of $\mathbb{E}[R_\alpha]$ from (6.7) is equal to 11.21% and 95% two sided confidence interval ranges from 15.91% to 6.51%. This gives us the cost of capital for Alpha and now it is a straightforward task to calculate their WACC.

$$\mathbb{E}[R_\alpha] = R_{risk-free} + \beta_\alpha(\mathbb{E}[R_{market}] - R_{risk-free}) \tag{6.7}$$

The formula for the WACC is given in equation (6.8). Refer to Table 4.2 for a description of all symbols in (6.8). All data for the WACC calculation, except for the cost of capital, can be found in the annual report of Alpha and we used the 2016 version. We found for Alpha that the average corporate tax rate (T_r) is estimated at 20%. The cost of debt (C_d) is calculated by dividing the interest expenses by total value of net debt, which gives an outcome of 6.35%. For the percentage of equity ($\frac{V_e}{MV_f}$) and percentage of debt ($\frac{V_d}{MV_f}$) we found respectively 61.27% and 38.73%.

$$\text{WACC} = \frac{V_e}{MV_f}C_e + \frac{V_d}{MV_f}C_d(1 - T_r) \tag{6.8}$$

The estimated value of the WACC is equal to 8.84% with a two sided 95% confidence interval ranging from 5.96% to 11.71%. The lower bound for the cost of capital is almost equal to the cost of debt, but this seems to be unrealistic. Equity investors would normally require a higher return than credit investors, because their risk of investment is higher. Nonetheless, we will use the estimates that we have found to reflect various plausible future scenarios.

Chapter 7

Numerical results

In this chapter we present the numerical results that we gained from solving the two-stage stochastic programming problem presented in Section 5.1. We use the SAA algorithm to solve this problem which is explained in detail in Section 5.2. This problem is a multi objective problem and we solved it for different weights of each objective.

7.1 Problem instances

Our stochastic programming objective function contains two objectives with the relative weighted lead time performance difference and the supply chain cost. We introduced λ as the variable to weigh both objectives in the total objective function. The lead time performance difference is multiplied by λ and the total cost is multiplied by $(1-\lambda)$. For λ its values are bounded between one and zero, i.e. $0 \geq \lambda \geq 1$. Setting λ equal to one or zero means we completely exclude one of two objectives and we want to avoid this. We vary λ between 0.1 and 0.9 with increments of 0.1 this leads to nine different possible values and thus nine different problem instances. To solve the problems we use the SAA algorithm, which requires three parameters to be set. The parameters for the SAA algorithm are M, N and N' and these respectively represent the number of samples and the number of scenarios for the optimization problem and for the estimation. We choose to let $M = 20$, because we observed in many sources that this is standard procedure for the SAA (Schütz et al., 2009; Santoso et al., 2005). For the number of scenarios for the optimization problem we choose for $N = 5, 10, 20$. The number of scenarios for the estimation of the "true" problem's objective value is set to $N' = 1000$. We want to choose a number of scenarios that is much bigger than N and we observed this was the standard procedure in literature (Schütz et al., 2009; Santoso et al., 2005). Table 7.1 depicts the results that we have gained from our solved problem instances. The relative optimality gap for the solution is defined by equation (7.1).

$$\text{relative gap}_{N,M,N'}(\bar{y},\lambda) := \frac{\tilde{f}_{N'}(\bar{y},\lambda) - \bar{v}_{N,M}}{\tilde{f}_{N'}(\bar{y},\lambda)} \tag{7.1}$$

We are also interested in the confidence interval associated with the relative optimality gap, see equation (7.2) for the two-sided confidence interval. Where $z_\alpha =: \Phi^{-1}(1-\alpha)$ and Φ denotes the cumulative density function of the standard normal distribution.

$$\text{confidence interval}_{N,M,N'}(\bar{y},\lambda) := \frac{\tilde{f}_{N'}(\bar{y},\lambda) - \bar{v}_{N,M} \pm z_\alpha(\sigma^2_{N'}(\bar{y},\lambda) + \sigma^2_{\bar{v}_{N,M}})^{\frac{1}{2}}}{\tilde{f}_{N'}(\bar{y},\lambda)} \tag{7.2}$$

We chose to calculate the 90% confidence interval for all optimality gaps. For every sample size we selected the solution that gave the lowest objective value and defined this as $\bar{y}$ in (7.1) and (7.2). We argue that the sample sizes could be further increased, but our biggest sample sizes gave results that are of good enough quality. Improving the quality is possible at the expense of more computational effort. We argue this gain is minimalistic when inspecting the relative optimality gaps for sample sizes of 20, because they are almost all below 1%.

Table 7.1: Optimality gaps from SAA algorithm in percentages

λ	N	*Relative optimality gap (%)*	*Half width confidence interval (%)*
0.1	5	3.53	2.9
	10	1.48	2.56
	20	1.08	2.92
0.2	5	3.35	2.9
	10	1.42	2.96
	20	0.85	3.18
0.3	5	3.9	2.86
	10	1.45	2.84
	20	0.63	2.58
0.4	5	3.49	3.17
	10	1.37	2.67
	20	0.61	3.18
0.5	5	3.62	3.18
	10	1.83	2.88
	20	0.58	2.74
0.6	5	3.76	3.33
	10	1.41	2.96
	20	0.65	3.16
0.7	5	3.48	3.38
	10	1.74	3.84
	20	1.0	3.32
0.8	5	3.62	5.22
	10	1.66	4.4
	20	0.93	4.46
0.9	5	3.71	17.28
	10	2.32	14.82
	20	-0.51	13.96

In Table 7.1 we see a negative value for an optimality gap. It is an estimation, of course, and that it is negative suggest that the approximation is close to zero. This is possible because we use random variables to calculate $\tilde{f}_{N'}(\bar{y}, \lambda)$ and $\bar{v}_{N,M}$. It is important that all confidence intervals contain an optimality gap greater than zero and this is the case for Table 7.1. Furthermore, it is quite remarkable that the optimality gaps confidence intervals are significantly bigger when $\lambda = 0.9$. In this scenario we have the lead time performance difference weighing nine times heavier than the supply chain cost. The lead time performance difference is a number below zero and therefore the objective function's value gets smaller, but the variance remains at the same absolute value for other values of λ. This causes the confidence intervals for the optimality gaps to become bigger for relative high values of λ.

7.2 Solution properties

Cumulative probability curves

We calculated the values of $\tilde{f}_{N'}(\bar{y}, \lambda)$ for various parameter settings. For the solution $\bar{y}$ we chose the solution that gave the lowest objective value with $N = 20$. We specifically look at the behavior of the two separate objectives. Figure 7.1 depicts how the cumulative density curves of the total cost are distributed for various value of λ. The red line represents Alpha's current SCND without making any changes. The most left line represents the case where $\lambda = 0.1$ and the more right the curves are located, the higher λ is. We created the same plot for the lead time performance difference, see Figure 7.2. In

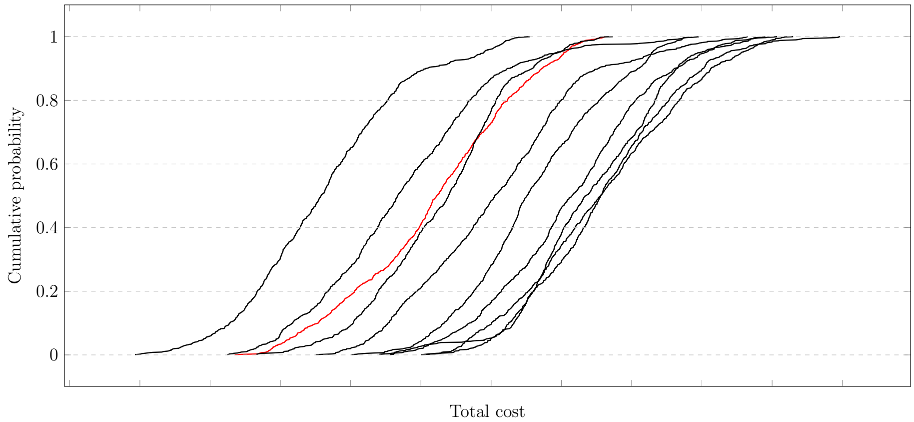

Figure 7.1: Cumulative probability curves of the total cost for different values of λ and N=20

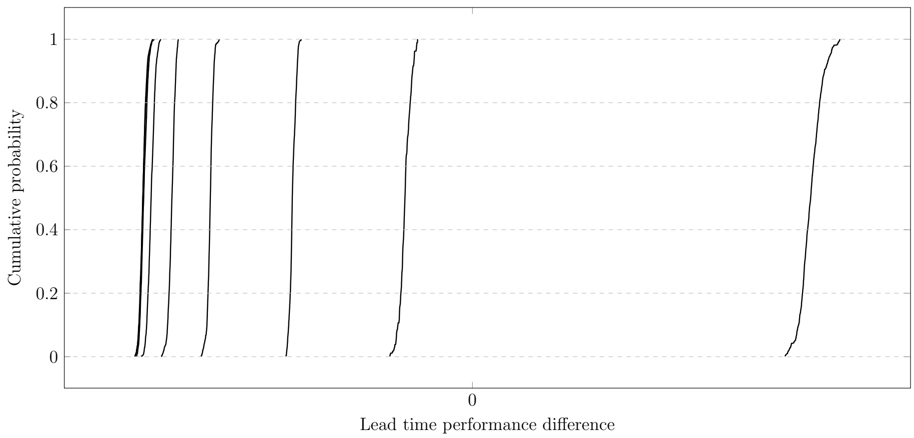

Figure 7.2: Cumulative probability curves of lead time performance difference for different values of λ and N=20

this plot the case where $\lambda = 0.1$ is the most right curve, which is the opposite of Figure 7.1. The differences between the curves values on the horizontal axis in Figure 7.2 are significantly bigger than in Figure 7.1.

Pareto front

We obtain a Pareto front from solving all problem instances in Table 7.1. For each value of λ in Table 7.1 we plotted the relative weighted lead time performance difference against the total cost. Figure 7.3 depicts the Pareto front from our problem instances where the white circles and the red triangle respectively represent all problems instances and Alpha's current SCND. We can draw an imaginary line through connecting all white

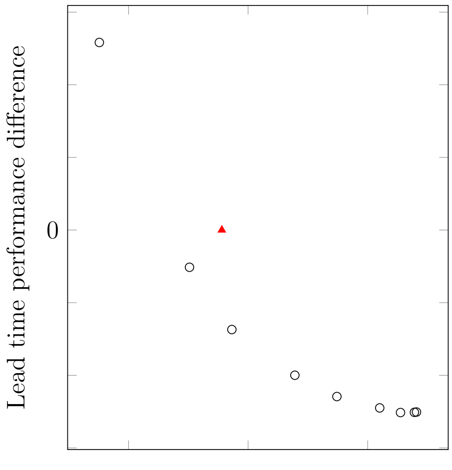

Figure 7.3: Pareto front from calculating the objective function with N'=1000

circles and this is the Pareto front for this particular bi-objective problem. All solutions left of the Pareto front are infeasible solutions and all solutions right of the Pareto front are the dominated solutions. It is clear that the Alpha's current SCND is dominated by the solutions from our problem instances. There even seems to be a SCND which has almost equal cost, but a negative relative weighted lead time performance difference this means it results in shorter lead times.

Assembly locations

The previous statistics and plots describe nothing about the distribution of assembly locations for the product families over Alpha's factories. In Table 7.2 we have an overview of the current situation for Alpha's factories. In the new situation we only have locations Jupiter, Saturn and Uranus as assembly locations, so Neptune is excluded. This does not mean that Alpha should close the Neptune location, because we only had useful data for five product families which are currently produced at the Neptune location. In Table

Table 7.2: Current situation for assembly locations

	Jupiter	*Saturn*	*Uranus*	*Neptune*
Number of locations	26	25	23	5

7.3 we summarized the results for our chosen values of λ and N. From the results in Table 7.3 we can see that if λ is relatively low, then most product families are moved to the Saturn location. The Saturn location has very competitive labor cost compared to the other factories. When the value of λ increases the number of product families for the Saturn factory decreases, because the lead time performance difference gets more important and most components are produced in either the Jupiter or Uranus factory.

Table 7.3: Assembly locations for different values of λ and sample size N

λ		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
N	*Location*									
5	*Jupiter*	1	8	15	26	29	32	35	35	35
	Saturn	58	49	41	29	25	22	19	19	19
	Uranus	20	22	23	24	25	25	25	25	25
10	*Jupiter*	2	8	14	26	29	32	35	35	35
	Saturn	56	48	41	27	24	21	18	18	18
	Uranus	21	23	24	26	26	26	26	26	26
20	*Jupiter*	2	8	14	27	30	32	35	35	35
	Saturn	56	48	41	27	23	21	18	18	18
	Uranus	21	23	24	25	26	26	26	26	26

7.3 Verification and validation

We describe the process of verification and validation of our proposed model. The verification process serves to check whether the model is build correctly, i.e. to check if it delivers the optimal solution. The process of validation is to determine whether the model represents the real system we are trying to model.

Verification

The verification of a model is to check whether the model behaves as it should. We want to be sure that this model gives the optimal outcome as a result when we use it. Our proposed model is stochastic programming model which is solved by using the SAA algorithm. The original stochastic programming problem boils down to a MILP because of the SAA algorithm. We want to minimize our objective function for the generated scenarios. The MILP problems are solved by using the Gurobi solver. To verify our model we create some extreme scenarios and observe whether the expected outcome is delivered by our model. In addition, we perform verification by using historical data and selecting some product families. From this data we build some case analysis and calculate the total cost and relative weighted lead time performance by using an Excel sheet. We check whether our model gives the same results as what we calculate in the Excel sheet.

Validation

A model is validated to ensure it represents the real-system sufficiently well. For this model we checked face validity, which means we checked the model's output with several people at Alpha who have crucial knowledge about the supply chain processes. After some iterations in the modeling phase we concluded that it gave reasonable results. Thus we have validated our model through face validation.

Chapter 8

Implementation

In this chapter we discuss how our decision support tool should be implemented at Alpha. First of all, we want to highlight that we created a quantitative model for a real system, namely Alpha's supply chain. A quantitative model will always be a simplified representation of the real world and its results should be interpreted with caution. That being said, our decision support tool aims to provide Alpha with accurate data to make fact based decisions about their new SCND.

Our model has been written in the Python language by using the PuLP package, this is a package that enables the user to write linear programming models in a straightforward manner. To solve the linear programming models we used the Gurobi solver, we chose this solver because it gave a free unlimited academic license and can handle MILP models. Our decision support tool is build in Microsoft Excel, because Alpha's decision makers are familiar with this software. In addition, it is quite cumbersome to install Python and all package dependencies on every computer to use this tool. Furthermore, they would have to install a solver which is for non-academic use and this can be a costly exercise. The Excel tool calculates the same objective functions, i.e. the cost and relative weighted lead time performance difference. Within the Excel tool they can specify their own scenarios for the future on a two year horizon. The choice for the fixed two year horizon is related to the fact that the forecast of demand for Alpha looks maximum two years ahead.

Our tool should be used as support in their decision making process. The lead time difference for the various scenarios can be used to calculate also the impact on the inventory levels of the finished goods. We set the scope of our project and left out these inventory levels, but the impact can be significant and should be calculated to estimate the impact on the total cost. It is obvious that safety stocks can be reduced if the lead times are shorter and inventory turns can grow. These are of course important performance indicators for the supply chain management and thus should not be ignored. The decision makers can combine the results from our report with their Excel tool to make decision about their SCND.

To build this tool we collected data from Alpha's various ERP systems and we made a snapshot of all combined historical data in our model. It would be beyond the scope of this project to link our model to all these systems, but this means we created a snapshot analysis of all available data up until this September 2017. The maintenance of this tool is therefore not necessary, but it can be performed by manually collecting missing future data.

Chapter 9

Conclusions and recommendations

This is the last chapter of this report and it shows what conclusions can be drawn from our research project. In addition, we provide directions for future research in the field of SCND and stochastic programming research.

9.1 Conclusions

In this thesis, we have considered a SCND problem which is focused on determining the right location for assembly of products. Currently, Alpha produces most of the components for the assembly operations at two locations, namely Jupiter and Uranus. Some years ago they made a decision to move some of their assembly operations to location Saturn, but this decision was solely based on the low labor cost related to this location. Now, Alpha wants to do an more thorough analysis on what the consequences are of this decisions and reconsider the locations for assembly operations. This project excludes the relocation of the production location for components. It is in particular aimed at the location of the assembly operations.

We have shown in this report that this problem can be classified as a SCND problem. Then we chose to solve this SCND problem with the stochastic programming approach. This approach is based on regular linear programming techniques, but it allows for the inclusion of the uncertainty of parameter values. In this research project, we tried to model the uncertain parameters by either a theoretical probability distribution or an empirical distribution. We chose to treat the transportation and assembly operations expenses as parameters with uncertainty regarding their exact value in the future. The difficulty of the stochastic programming approach is the huge number of possible scenarios when combining all the random variables in one model. To solve this problem, we have used the SAA algorithm, which is a technique that basically samples a number of scenarios. Furthermore, this technique shows that stochastic programming problem reduces to a MILP problem. It is well known that these problems can be solved straight-forward by using a MILP solver.

The stochastic programming approach we use is somewhat different from what we found in existing literature. We proposed a bi-objective problem where each objective is assigned a weight. The objective we proposed in our model are the total cost which is derived from the DRCF model and the relative weighted lead time performance difference. The objectives' weights were normalized which can be done without loss of generality, because there are only two objectives. We solved the problem for various scenarios for the parameter values by using the SAA algorithm. Furthermore, we also varied the weights for the two objectives. This illustrates the trade off between the lead time performance and associated cost. We are now able to answer to earlier formulated main research question. The main research question of our research project is reformulated below.

How can Alpha create more value looking at their production locations for their complete product portfolio. Specifically, do they have to change some production/assembly operation locations within their supply chain?

We analyzed Alpha's complete product portfolio in this research project. First, we made a selection of products that have some useful available data to do such analysis. We implemented the selection of products in our stochastic programming model and this gave some interesting insights. Alpha's current SCND is dominated by a solution we found. There is a difference of approximately 0.6% between the present value of the total cost of respectively the current and the dominating configuration. This difference is for the complete two year horizon. Where the dominating configuration has lower total cost and a better lead time performance than the current configuration. In addition, our results show that Neptune should not be included the SCND. We only included a small selection of the Neptune products in our model. For the Neptune location there was not much useful data available to make a proper analysis. So this does not mean that the Neptune location should closed according to our results. In the next section we discuss the recommendations following from our results in this project.

9.2 Recommendations

We observe that Alpha's current SCND for the selection of products is dominated in terms of total cost and lead time performance by other alternatives. In Chapter 7 we included a plot of a Pareto front based on our solutions for the stochastic programming problem. Based on our model we conclude that Alpha can reduce the total cost and increase its lead time performance if it changes some of its assembly locations. In addition, our results show there exist even more options for changing the assembly locations. First, Alpha could decide to further increase its lead time performance at the expense of a higher total cost than the current situation. Contrary, they could also choose to reduce the total cost compared to the current situation, but the lead time performance will decrease according to our model.

The analysis for Alpha's SCND can be further extended by including the lead time impact on stock levels for finished goods in the hubs. In our analysis we have calculated the total cost related to the production and transportation of components, the assembly of the components and transportation of finished goods to the hubs locations. We did calculate the lead time performance difference and this information can be used to estimate the impact on the stock levels of finished goods in all of their hubs.

References

Birge, J. R., & Louveaux, F. (2011). *Introduction to stochastic programming*. Springer Science & Business Media.

Brealey, R. A., Myers, S. C., Allen, F., & Mohanty, P. (2012). *Principles of corporate finance*. Tata McGraw-Hill Education.

Fama, E. F., & French, K. R. (2004). The capital asset pricing model: Theory and evidence. *The Journal of Economic Perspectives*, *18*(3), 25–46.

Kleywegt, A. J., Shapiro, A., & Homem-de Mello, T. (2002). The sample average approximation method for stochastic discrete optimization. *SIAM Journal on Optimization*, *12*(2), 479–502.

Klibi, W., Martel, A., & Guitouni, A. (2010). The design of robust value-creating supply chain networks: a critical review. *European Journal of Operational Research*, *203*(2), 283–293.

Kouvelis, P., & Yu, G. (1997). *Robust discrete optimization and its applications* (Vol. 14). Springer US.

Mak, W.-K., Morton, D. P., & Wood, R. K. (1999). Monte carlo bounding techniques for determining solution quality in stochastic programs. *Operations research letters*, *24*(1), 47–56.

Melo, M. T., Nickel, S., & Saldanha-Da-Gama, F. (2009). Facility location and supply chain management–a review. *European Journal of Operational Research*, *196*(2), 401–412.

Mitroff, I. I., Betz, F., Pondy, L. R., & Sagasti, F. (1974). On managing science in the systems age: two schemas for the study of science as a whole systems phenomenon. *Interfaces*, *4*(3), 46–58.

Norkin, V. I., Pflug, G. C., & Ruszczyński, A. (1998). A branch and bound method for stochastic global optimization. *Mathematical programming*, *83*(1-3), 425–450.

Revelle, C. S., Eiselt, H. A., & Daskin, M. S. (2008). A bibliography for some fundamental problem categories in discrete location science. *European Journal of Operational Research*, *184*(3), 817–848.

Rosenhead, J., Elton, M., & Gupta, S. K. (1972). Robustness and optimality as criteria for strategic decisions. *Operational Research Quarterly*, *23*(4), 413–431.

Royal Dutch Shell Group. (2005). *Shell global scenarios to 2025: The future business environment: Trends, trade-offs and choices.*

Santoso, T., Ahmed, S., Goetschalckx, M., & Shapiro, A. (2005). A stochastic programming approach for supply chain network design under uncertainty. *European Journal of Operational Research*, *167*(1), 96–115.

Schütz, P., Tomasgard, A., & Ahmed, S. (2009). Supply chain design under uncertainty using sample average approximation and dual decomposition. *European Journal of Operational Research*, *199*(2), 409–419.

Shapiro, A. (2007). Stochastic programming approach to optimization under uncertainty. *Mathematical Programming*, *112*(1), 183–220.

Stadtler, H., & Kilger, C. (2008). Supply chain management and advanced planning. *Concepts, Models, Software and Case Studies*, *4*.

Stewart, T. J. (2005). Dealing with uncertainties in MCDA. In J. Figuera, S. Greco, & M. Ehrgott (Eds.), *Multiple criteria decision analysis: state of the art surveys* (chap. 11). Springer International Series.

van Aken, J., Berends, H., & Van der Bij, H. (2012). *Problem solving in organizations: A methodological handbook for business and management students.* Cambridge University Press.

Will M. Bertrand, J., & Fransoo, J. C. (2002). Operations management research methodologies using quantitative modeling. *International Journal of Operations & Production Management*, *22*(2), 241–264.

Appendices

Appendix A

Cullen and Frey graph

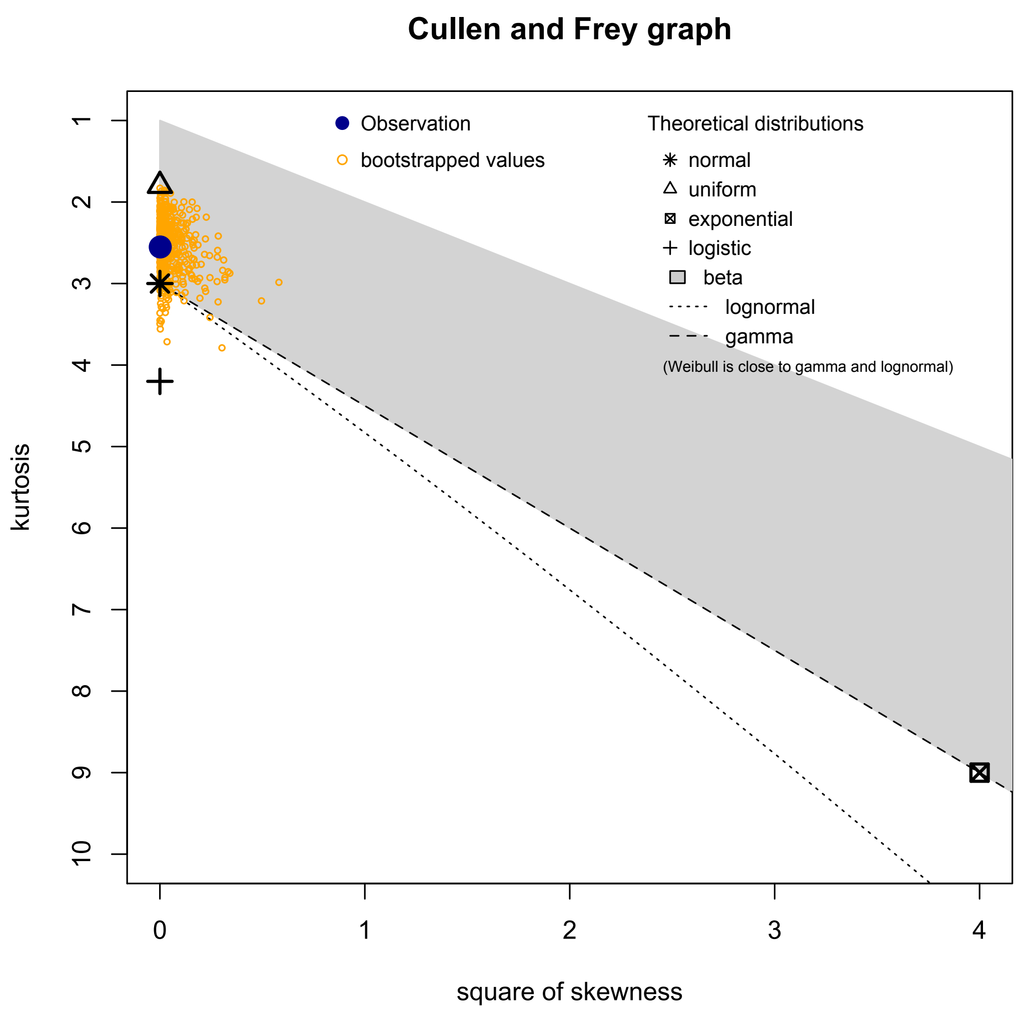

Figure A.1: Cullen and Frey graph